

# Begynnende lese- og staveutvikling

*En analyse av longitudinelle data med  
autoregressive modeller og latente  
vekstkurvemodeller*

Hilde Olsen



Masteroppgave i utdanningsvitenskapelig metodologi  
Det utdanningsvitenskapelige fakultet

UNIVERSITETET I OSLO

14.06.2011





# **Begynnende lese- og staveutvikling**

*En analyse av longitudinelle data med autoregressive modeller og latente vekstkurvmodeller*

© Hilde Olsen

2011

Begynnende lese- og staveutvikling. En analyse av longitudinelle data med autoregressive modeller og latente vekstkurvmodeller

Hilde Olsen

<http://www.duo.uio.no/>

Trykk: Nilz Trykkerier, Oslo

# Sammendrag

Mange teoretiske hypoteser fokuserer på forandring, og man må ha longitudinelle data for fullt ut å kunne evaluere slike hypoteser (Bollen & Curran, 2006). Jeg analyserer deler av data fra en større longitudinell studie av den begynnende lese- og staveutviklingen hos et representativt utvalg norske barn. Mine data er repeterte målinger av nonordlesing, nonordstaving, ortografisk lesing og ordstaving fra andre til fjerde klasse. Litt over 200 barn er målt fire ganger på alle lese- og stavevariablene: I november/desember i andre klasse (rundt tre måneder etter oppstart av formell lese- og staveundervisning), i mai/juni i andre klasse, i november/desember i tredje klasse og i november/desember i fjerde klasse. Jeg har også med målinger av to kognitive variabler (fonologisk bevissthet og alfanumerisk RAN). Disse er målt kun i november/desember i andre klasse.

Samspeilet mellom begynnende lese- og staveutvikling har i liten grad blitt systematisk undersøkt og det har heller ikke vært gjennomført så mange longitudinelle studier på dette feltet (Leppänen, Nieme, Aunola & Nurmi, 2006). Det eksisterer ulike teorier om hvordan forholdet mellom lese- og staveutviklingen er: I følge Frith (1985), bytter lesing og staving på å influere hverandre i bestemte faser av utviklingen. I følge Ehri (1995, 2000) er det å lære seg å lese og stave ord tett knyttet sammen. Lesing og staving utvikler seg sammen og påvirker hverandre gjensidig; lesingen påvirker stavingen, men på samme tid påvirker stavingen lesingen.

I datamaterialet som ligger til grunn for analysene i denne oppgaven, er lesing og staving målt gjennom lesing og staving av enkeltord, både nonord og konvensjonelle ord. Det er altså snakk om måling av den mer ”tekniske” siden, ikke forståelsesaspektet, av begynnende lese- og staveferdigheter, nærmere bestemt ordavkodingsferdigheter. Lesing og staving av nonord regnes for å være et godt mål på tidlige ordavkodingsferdigheter, og kan sies å være mer ”fonologiske i sin natur” enn ortografisk lesing og konvensjonell ordstaving.

Jeg stiller følgende forskningsspørsmål: Er det nonordstavingen som forklarer nonordlesingen tidlig i utviklingen og den ortografiske lesingen som forklarer ordstavingen litt senere i utviklingen, eller er det slik at lesing og staving utvikler seg sammen og påvirker hverandre gjensidig? Disse hypotesene testes ut både med og uten fonologisk bevissthet og alfanumerisk RAN som kontrollvariabler (kovariater). Det andre spørsmålet som stilles er hvor godt

autoregressive modeller og latente vekstkurvemodeller egner seg til å si noe om den gjensidige utviklingen av begynnede lese- og staveferdigheter og forholdet mellom disse.

Autoregressive modeller og latente vekstkurvemodeller har ulike tilnærminger til analyse av repeterte målinger. Kort sagt kan man si at autoregressive modeller analyserer endringer i rangering over tid (er det de samme som er gode til å lese og stave på tidspunkt 1 og 2?), mens latente vekstkurvemodeller fokuserer på det enkelte individs faktiske vekst. Analysene er gjennomført ved hjelp av programmet Mplus, versjon 6.1 og SPSS, versjon 16.0.

Mine resultater gir ingen støtte til Friths teori om at det er nonordstaving som forklarer nonordlesing tidlig i utviklingen. Videre viser mine resultater at fonologisk bevissthet og RAN i stor grad forklarer den tidlige sammenhengen mellom nonordlesing og nonordstaving. Disse kovariatene forklarer også en del av sammenhengen mellom ortografisk lesing og ordstaving på det første måletidspunktet, men ikke fullt ut slik som for nonordlesing- og staving. En sentral innvending er at jeg ikke har data på lesing og staving målt i første klasse.

Mine analyser av ortografisk lesing og ordstaving gir noe, om enn ganske svak, støtte til Friths teori om at det er ortografisk lesing som forklarer staveutviklingen når barna har nådd et visst nivå av kjennskap til sammenhengen mellom grafemer og fonemer. Resultater fra den autoregressive modellen viser at ortografisk lesing målt på slutten av andre klasse predikerer ordstaving i midten av fjerde klasse når det kontrolleres for autoregressive effekter av staving på tidligere tidspunkt.

Vekstkurvemodellen indikerer at det er en gjensidig påvirkning mellom ortografisk lesing og ordstaving hele veien, og gir dermed støtte til Ehrlis teori. Her er det en sterk sammenheng mellom den uforklarte variansen til de to slopefaktorene (altså de vekstfaktorene som representerer den lineære veksten). Dette kan tolkes som at den lineære veksten i ortografisk lesing og ordstaving har noe til felles utover det som forklares av de respektive ferdighetene på første måletidspunkt og kovariatene fonologisk bevissthet og alfanumerisk RAN.

Vekstkurvemodellene gir mer detaljert informasjon om vekst enn de autoregressive modellene ved at de sier noe om sammenhengen mellom hvor man starter og hvordan man vokser. Vekstkurvemodellene gjør det også mulig å modellere ikke-lineær vekst. Mine resultater, der jeg finner at veksten er ikke-lineær for alle de fire lese- og stavevariablene, synliggjør også

viktigheten av å måle lenge nok når man skal si noe om utviklingen av denne typen ferdigheter.

I min undersøkelse ser det ut til at de latente vekstkurvmodellene ikke fanger opp nyansene i relasjonen mellom lesing og staving. Dersom Friths (1985) teori stemmer, er det snakk om tidsspesifikke relasjoner, ikke relasjoner der veksten er generalisert over alle måletidspunktene. I den autoregressive modellen predikerer ortografisk lesing målt i slutten av andre klasse ordstaving målt i midten av fjerde klasse. Denne sammenhengen finner jeg ikke i vekstkurvmodellen. I vekstkurvmodellen fanges heller ikke sammenhengen mellom fonologisk bevissthet målt i starten av andre klasse og nonordstaving målt i slutten av andre klasse opp. Dette fordi fonologisk bevissthet har størst betydning for nonordstaving helt i starten av utviklingen. Denne sammenhengen forsvinner i en vekstkurvmodell der veksten generaliseres over så mange år som den gjør i min undersøkelse. Det ser altså ut til at de autoregressive modellene fanger opp tidsspesifikke relasjoner bedre enn vekstkurvmodellene.

De statistiske analysemetodene utvikles stadig og blir mer og mer sofistikerte og fleksible med mulighet for kompleks modellering. Det er alltid en fare for at denne kompleksiteten går ut over tolkbarheten. Det er uansett umulig å bevise at en modell er korrekt gjennom statistiske analyser.







# Forord

Hva er det som har fått meg til å velge et studium (og dermed også en masteroppgave) der gjennomføring og evaluering av kvantitative analyser og metodiske forhold i utdanningsvitenskapen er i fokus?

For det første er det ikke tilfeldig at studiet er innenfor utdanningsvitenskapen: Jeg har mange års pedagogikkstudier bak meg og jeg har jobbet administrativt innenfor disse fagfeltene i hele mitt yrkesliv; de siste årene som forskningskonsulent ved Institutt for spesialpedagogikk. Årene som forskningskonsulent har økt min interesse for forskning, og dessuten hjelper det jo å være glad i matematikk og statistikk. Kombinasjonen av min interesse for pedagogikk og spesialpedagogikk og det kvantitative er rett og slett en uimotståelig kombinasjon. Det å skrive en masteroppgave har vært en lærerik og artig prosess, og jeg føler meg utrolig privilegert som har hatt muligheten til å gjennomføre dette.

Jeg har hatt ulønnet permisjon fra jobben som forskningskonsulent ved Institutt for spesialpedagogikk for å gjennomføre masterstudiet. Takk til instituttet som var villige til å gå inn på en slik ordning. Min kjære Torgrim fortjener en varm takk for å ha støttet meg så helhjertet hele veien. Til slutt må jeg rette en stor takk til min faglig sterke og konstruktive veileder Arne Lervåg. Arne har dessuten sjenerøst gitt meg tilgang til sine egne data fra en større longitudinell undersøkelse om begynnende lese- og staveutvikling. Det er derfor flere grunner til å si at uten Arne hadde denne masteroppgaven aldri blitt noe av. Tusen takk!

Oslo, juni 2011

Hilde Olsen



# Innholdsfortegnelse

1	Innledning.....	1
1.1	Bakgrunn for valg av tema .....	1
1.2	Oppbygging av oppgaven.....	4
2	Begynnende lese- og staveutvikling.....	5
2.1	Teorier om lese- og staveutvikling .....	5
2.2	Variabler som kan forklare begynnende lese- og staveutvikling .....	8
2.2.1	Fonologisk bevissthet.....	8
2.2.2	Rapid Automatized Naming (RAN).....	10
2.3	Variasjoner mellom språk.....	12
2.4	Tidligere forskning på forholdet mellom begynnende lese- og staveutvikling .....	13
2.5	Forskningsspørsmål .....	14
3	Metode.....	16
3.1	Strukturelle likningsmodeller .....	16
3.1.1	Latente variabler.....	17
3.1.2	Estimering av parametre i modellen.....	18
3.1.3	Modelltilpasning.....	19
3.1.4	Evaluerings av modeller .....	22
3.1.5	Manglende data .....	23
3.2	Latente vekstkurvmodeller .....	25
3.2.1	Lineær latent vekstkurvmodell .....	26
3.2.2	Ikke-lineære latente vekstkurvmodeller .....	29
3.2.3	Bivariat latent vekstkurvmodell .....	30
3.3	Autoregressive modeller.....	31
3.3.1	Univariat latent autoregressiv modell .....	31
3.3.2	Bivariat latent autoregressiv modell.....	33
3.4	Data.....	34
3.4.1	Deltakere .....	34
3.4.2	Lese- og skriveopplæring .....	35
3.4.3	Design og prosedyre.....	35
3.4.4	Tester.....	36
4	Resultater.....	38

4.1	Deskriptiv statistikk/korrelasjoner.....	38
4.2	Målemodeller.....	41
4.3	Latente vekstkurvmodeller .....	41
4.3.1	Ubetingede univariate latente vekstkurvmodeller .....	42
4.3.2	Bivariate latente vekstkurvmodeller .....	50
4.4	Autoregressive modeller.....	58
4.4.1	Univariate latente autoregressive modeller .....	58
4.4.2	Bivariate autoregressive modeller.....	60
5	Diskusjon.....	64
5.1	Lesing og staving av nonord.....	64
5.2	Ortografisk lesing og ordstaving .....	65
5.3	Forholdet mellom begynnende lese- og staveutvikling .....	67
5.4	Latente vekstkurvmodeller .....	68
5.5	Autoregressive modeller.....	70
5.6	Alternative modeller .....	71
5.7	Hvorfor ikke se seg blind på modelltilpasningen .....	72
5.8	Avsluttende kommentarer.....	72
	Litteraturliste .....	75
	Vedlegg .....	79

## Oversikt over figurer og tabeller

<b>Figur 1</b>	Friths sekstrinnsmodell for tilegnelsen av lese- og staveferdigheter.....	6
<b>Figur 2</b>	Ubetinget, lineær latent vekstkurvmodell for fire måletidspunkter .....	27
<b>Figur 3</b>	Univariat latent autoregressiv modell med målemodell .....	32
<b>Figur 4</b>	Bivariat latent autoregressiv modell med kryssladinger (uten målemodellen) .....	34
<b>Tabell 1</b>	Deskriptiv statistikk og korrelasjoner for observerte kognitive variabler T2 .....	39
<b>Tabell 2</b>	Deskriptiv statistikk og korrelasjoner for ortografisk lesing og nonordlesing T2-T5 .....	39
<b>Tabell 3</b>	Deskriptiv statistikk og korrelasjoner for ordstaving og nonordstaving T2-T5 .....	40
<b>Figur 5</b>	Målemodell for latente kovariater T2 .....	41
<b>Figur 6</b>	Observert og estimert gjennomsnitt for nonordlesing, lineær modell .....	42
<b>Figur 7</b>	Observert og estimert gjennomsnitt for nonordlesing, kvadratisk modell .....	43
<b>Figur 8</b>	Observert og estimert gjennomsnitt for nonordlesing, modell med to fritt estimerte slopefaktorer.....	44
<b>Tabell 4</b>	Modelltilpasning for ubetingede latente vekstkurvmodeller .....	45

<b>Tabell 5</b> Ubetingede latente vekstkurvmodeller for ortografisk lesing, ordstaving, nonordstaving og nonordlesing .....	46
<b>Figur 9</b> Kvadratisk latent vekstkurvmodell .....	47
<b>Figur 10</b> Estimerte gjennomsnitt for ortografisk lesing, kvadratisk vekstkurvmodell .....	48
<b>Figur 11</b> Estimerte gjennomsnitt for ordstaving, kvadratisk vekstkurvmodell .....	49
<b>Figur 12a</b> Bivariat latent vekstkurvmodell nonordlesing og nonordstaving uten kryssregresjoner .....	52
<b>Figur 12b</b> Bivariat latent vekstkurvmodell nonordlesing og nonordstaving med kryssregresjoner .....	53
<b>Figur 12c</b> Bivariat latent vekstkurvmodell nonordlesing og nonordstaving med kovariater .....	54
<b>Figur 13a</b> Bivariat latent vekstkurvmodell ortografisk lesing og ordstaving uten kryssregresjoner .....	56
<b>Figur 13b</b> Bivariat latent vekstkurvmodell ortografisk lesing og ordstaving med kovariater .....	57
<b>Figur 14</b> Univariat latent autoregressiv modell med målemodell .....	58
<b>Tabell 6</b> Estimerte univariate latente autoregressive modeller.....	59
<b>Tabell 7</b> Modelltilpasning univariate latente autoregressive modeller.....	60
<b>Figur 15a</b> Bivariat latent autoregressiv modell nonordlesing og nonordstaving .....	61
<b>Figur 15b</b> Bivariat latent autoregressiv modell nonordlesing og nonordstaving med kovariater .....	61
<b>Figur 16a</b> Bivariat latent autoregressiv modell ortografisk lesing og ordstaving .....	62
<b>Figur 16b</b> Bivariat latent autoregressiv modell ortografisk lesing og ordstaving med kovariater.....	63





# 1 Innledning

## 1.1 Bakgrunn for valg av tema

Arne Lervåg ved Pedagogisk forskningsinstitutt har gjennomført en omfattende longitudinell, korrelasjonell studie av begynnende lese- og staveutvikling. Jeg har vært så heldig å få låne deler av data fra dette prosjektet til bruk i min masteroppgave. I Lervågs studie ble 228 elever testet 5 ganger fra høsten i første klasse (før oppstart av formell lese- og staveopplæring) til høsten i fjerde klasse. Dataene er fra perioden med læreplanen L97 (1997-2006) der man ikke startet med formell lese- og staveopplæring før i andre klasse. Jeg benytter meg ikke av de kognitive variablene som ble målt i første klasse. Mine data er repeterte målinger av nonordlesing, nonordstaving, ortografisk lesing og ordstaving fra andre til fjerde klasse. Litt over 200 barn er målt fire ganger på alle lese- og stavevariablene: I november/desember i andre klasse (rundt tre måneder etter oppstart av formell lese- og staveundervisning), i mai/juni i andre klasse, i november/desember i tredje klasse og i november/desember i fjerde klasse. Jeg har også med målinger av to kognitive variabler (fonologisk bevissthet og alfanumerisk RAN). Disse er bare målt i november/desember i andre klasse.

Hovedfokuset i oppgaven vil være på kvantitative analyser og metodiske forhold, men det er heller ikke tilfeldig at jeg ønsker å gjøre noe innenfor lese- og skrivefeltet. Å kunne være en funksjonell leser og skriver er helt grunnleggende både for å kunne delta i samfunnet og for å tilegne seg annen kunnskap. Å vite mer om utviklingen av disse ferdighetene og hva som predikerer denne utviklingen har betydning for hvordan man driver grunnleggende lese- og skriveopplæring. For å forstå hva lese- og skrivevansker er, må en forstå hvordan lesing og skriving normalt utvikler seg. Hvis ikke, blir det vanskelig å identifisere hvem de svake leserne er, og også vanskelig å sette inn relevante tiltak. Det er viktig med kunnskap om dette for å kunne fange opp barn som er i faresonen for å utvikle lese- og skrivevansker så tidlig som mulig.

Mye forskning har blitt gjort på leseutvikling. Det er ikke gjort like mye på staveutvikling. Samspillet mellom begynnende lese- og staveutvikling har i liten grad blitt systematisk undersøkt, og det har heller ikke vært gjennomført så mange longitudinelle studier på dette

feltet (Leppänen, Nieme, Aunola & Nurmi, 2006). Det eksisterer ulike teorier om hvordan forholdet mellom lese- og staveutviklingen er. Ifølge Frith (1985) bytter lesing og staving på å influere hverandre i bestemte faser av utviklingen. Ehri (1995, 2000) mener at det å lære seg å lese og stave ord er tett knyttet sammen. Lesing og staving utvikler seg sammen og påvirker hverandre gjensidig; lesingen påvirker stavingen, men på samme tid påvirker stavingen lesingen.

Mye av forskningen om begynnende lese- og staveutvikling er utført innenfor en inkonsistent ortografi som engelsk, selv om det i løpet av det siste tiåret også er publisert en del studier på dette feltet innenfor mer konsistente ortografier som tysk (Landerl & Wimmer, 2008), finsk (Leppänen et al., 2006), nederlandsk (de Jong & van der Leij, 1999) og norsk (Lervåg, Bråten & Hulme, 2009; Lervåg & Hulme, 2009; Lervåg & Hulme, 2010). Det er viktig å tolke forskningsresultatene i lys av hvilken ortografi studien er foretatt innenfor, fordi det er variasjoner mellom ortografier når det gjelder hvor ortografisk konsistente eller inkonsistente de er. Dette påvirker både hvor raskt og effektivt man utvikler lese- og staveferdigheter, og i hvilke perioder ulike kognitive og lingvistiske variabler bidrar til å forklare denne utviklingen (Lervåg & Hulme, 2010).

Det er enighet i forskningen om at fonologisk bevissthet er en viktig prediktor for utviklingen av begynnende lese- og staveferdigheter (Lervåg, Bråten & Hulme, 2009; de Jong & van der Leij, 1999; Leppänen et al., 2006; Landerl & Wimmer, 2008). Rapid Automatized Naming (RAN) ser ut til å ha betydning for leseflyt også utover det helt tidlige utviklingsstadiet (Verhagen, Aarnoutse & van Leeuwe, 2008; Lervåg & Hulme, 2009). Når det gjelder staving, er det er ikke så mange studier som har sett på betydningen av RAN, og disse studiene gir heller ikke et entydig bilde (Lervåg & Hulme, 2010).

Mange teoretiske hypoteser fokuserer på forandring, og man må ha longitudinelle data for fullt ut å kunne evaluere slike hypoteser (Bollen & Curran, 2006). Longitudinelle studier er bedre egnet enn tverrsnittstudier til å kunne si noe om retningen når det gjelder mulige årsaksforhold. Hvis fonologisk bevissthet målt på tidspunkt 1 korrelerer med lesing målt på tidspunkt 2, kan vi ikke trekke den slutningen at gode leseferdigheter på tidspunkt 2 er årsak til god fonologisk bevissthet på tidspunkt 1. Vi kan derimot trekke den slutningen at fonologisk bevissthet på tidspunkt 1 er en *mulig* årsak til senere leseferdigheter. Det er allikevel viktig å være klar over at sammenhengen mellom fonologisk bevissthet og lesing

også kan skyldes en variabel som ikke er målt, som for eksempel generell intelligens. (Hulme & Snowling, 2009).

Mange av studiene som har sett på effekten av ulike forklaringsvariabler på lesing og staving, har ikke vært longitudinelle, og har dermed ikke tatt hensyn til mulige autoregressive effekter (det vil si effekten av lesing og staving målt på et tidligere tidspunkt). Det er spesielt viktig å kontrollere for autoregressive effekter når man er interessert i utvikling. Dersom man ikke kontrollerer for lesing og staving på et tidligere tidspunkt, kan man ikke vite om det virkelig er en effekt av for eksempel fonologisk bevissthet på lesing og staving, eller om denne effekten heller skyldes individuelle forskjeller i lesing og staving på et tidligere tidspunkt (de Jong & van der Leij, 1999).

Jeg vil i oppgaven benytte meg av to ulike analysemetoder for repeterte målinger; latente vekstkurvmodeller og autoregressive modeller. Begge disse analysemetodene er applikasjoner av Structural Equation Modeling, SEM (strukturelle likningsmodeller). Latente vekstkurvmodeller og autoregressive modeller har derfor de samme fordelene som SEM. Noen av disse fordelene er

1. muligheten til å benytte maximum likelihood teknikker for manglende data
2. å kunne vurdere modelltilpasning og å kunne identifisere mulige kilder til dårlig modelltilpasning
3. muligheten til å inkludere latente kovariater og latente repeterte variabler i modellene (Bollen & Curran, 2006).

Når man benytter latente variabler, blir den variansen indikatorene har felles, operasjonalisert som sann skåre varians, som blir “overført til” den latente variabelen (Brown, 2006). Dette medfører at estimater av sammenhengen mellom begreper man ønsker å få vite noe om, blir mer nøyaktige.

Autoregressive modeller og latente vekstkurvmodeller har ulike tilnærminger til analyse av repeterte målinger. Kort sagt kan man si at autoregressive modeller analyserer endringer i rangering over tid (er det de samme som er gode til å lese/stave på måletidspunkt 1 og 2?), mens vekstkurvmodeller fokuserer på det enkelte individs faktiske vekst. I latente vekstkurvmodeller kan både gjennomsnittlig vekst og individuelle variasjoner modelleres i

den samme analysen (Parrila, Aunola, Leskinen, Nurmi & Kirby, 2005; Preacher, Wichman, MacCallum & Briggs, 2008).

## 1.2 Oppbygging av oppgaven

**I kapittel 2** redegjør jeg for begynnende lese- og staveutvikling og ulike variabler som kan bidra til å forklare denne utviklingen, samt forholdet mellom begynnende lese- og staveutvikling. Jeg tar også opp variasjoner mellom språk når det gjelder hvor ortografisk konsistente eller inkonsistente de er.

**I kapittel 3** redegjør jeg for valgte analysemetoder og andre metodiske forhold. Fordi både latente vekstkurvmodeller og autoregressive modeller er applikasjoner av Structural Equation Modeling, SEM (strukturelle likningsmodeller) tar jeg først for meg temaer som latente variabler, modelltilpasning, evaluering av modeller, estimering av parametre i modellen og manglende data. Med dette som bakgrunn, redegjør jeg mer spesifikt for latente vekstkurvmodeller og autoregressive modeller. Jeg redegjør også for deltakere, variabler og tester i undersøkelsen.

**I kapittel 4** presenteres resultatene av mine analyser. Jeg starter med deskriptiv statistikk og korrelasjoner og fortsetter med analyser av latente vekstkurvmodeller og latente autoregressive modeller. I de bivariate modellene (både vekstkurver og autoregressive) analyseres først nonordlesing og nonordstaving sammen og deretter ortografisk lesing og ordstaving.

**I kapittel 5** diskuteres resultatene i lys av både substansiell og metodisk teori.

## 2 Begynnende lese- og staveutvikling

Det er utviklingen av begynnende lese- og staveferdigheter som er tema for denne oppgaven. Lese- og staveferdigheter er målt gjennom lesing og staving av enkeltord, både nonord og konvensjonelle ord. Enkeltordene står dermed ikke i en sammenheng der man kan dra nytte av tekstsammenhengen til å identifisere ordene. Det er her snakk om måling av den mer ”tekniske” siden, og ikke forståelsesaspektet, av de begynnende lese- og staveferdighetene, nærmere bestemt ordavkodingsferdigheter.

Lesing og staving av nonord regnes for å være et godt mål på tidlige ordavkodingsferdigheter. I slike tester kan man ikke støtte seg på at man kjenner ordet fra før. For lesingen sin del innebærer ordavkodning å forstå hvilke lyder (fonemer) som representerer ulike bokstavkonstellasjoner (grafemer). For den begynnende stavingen handler det om å forstå hvilke bokstavkonstellasjoner (grafemer) som hører til hvilke lyder (fonemer). ” I tillegg til å lære bokstav-lyd-forbindelsene skal barna lære seg at noen lyder skrives på flere måter.” (Lyster, 1998, s. 46). Etterhvert som barna får praktisert lesing og konvensjonell staving av ord, lærer de at ord består av større enheter og bokstavkonstellasjoner som man finner igjen i ulike ord (Ehri, 1995).

### 2.1 Teorier om lese- og staveutvikling

Mange forskere har foreslått stadie- eller faseteorier for å beskrive hvordan prosesser utvikler og forandrer seg etterhvert som barn tilegner seg lese- og staveferdigheter. Utviklingsnivåene har forskjellige navn i de ulike teoriene, men de underliggende distinksjonene er de samme (Ehri, 2000). Jeg vil her redegjøre nærmere for to av disse teoriene: Friths stadiemodell og Ehri teori om ”sight word learning”.

Stadiemodellen til Frith (1985) deler utviklingen av leseferdigheter inn i tre faser som identifiseres av tre ulike strategier:

**Fase 1 - logografisk strategi:** Karakteriseres av umiddelbar gjenkjenning av kjente ord. Fremtredende grafiske trekk kan fungere som viktige ledetråder i denne prosessen. Rekkefølgen på bokstavene blir ofte ignorert, og fonologiske faktorer har ikke noe å si.

**Fase 2 - alfabetisk strategi:** Karakteriseres av kjennskap til og bruk av individuelle grafemer og fonemer og sammenhengen mellom disse (fonem er den minste lydenheten i språket som kan forandre et ords mening, og grafem er fonemets skriftlige uttrykk). Rekkefølgen på bokstavene og fonologiske faktorer spiller en avgjørende rolle.

**Fase 3 - ortografisk strategi:** I denne fasen skjer det en umiddelbar analyse av ord i ortografiske enheter, som for eksempel morfemer (språkets minste betydningsbærende enhet), uten fonologisk omkodning. Denne strategien skiller seg fra den logografiske ved at det er snakk om en systematisk analyse og ved at den er ikke-visuell. Den skiller seg fra den alfabetiske strategien ved at den tar i bruk større enheter og ved at den er ikke-fonologisk.

Frith (ibid.) mener videre at man kan dele literacy (lese- og staveferdigheter) inn i to komponenter: input (lesing) og output (staving), og at lesing og staving utvikler seg litt forskjellig. Med bakgrunn i dette, utvider Frith sin trefasemodell ved å dele hver fase inn i to trinn. Lesing og staving bytter på å være ”driveren” i utviklingsprosessen. Lesing er ”driveren” for den logografiske strategien, staving for den alfabetiske og lesing for den ortografiske. Denne utvidede modellen for tilegnelsen av lese- og staveferdigheter er illustrert i figur 1.

Trinn	Lesing	Staving
1a	Logografisk <sub>1</sub>	(symbolsk)
1b	Logografisk <sub>2</sub>	Logografisk <sub>2</sub>
2a	Logografisk <sub>3</sub>	Alfabetisk <sub>1</sub>
2b	Alfabetisk <sub>2</sub>	Alfabetisk <sub>2</sub>
3a	Ortografisk <sub>1</sub>	Alfabetisk <sub>3</sub>
3b	Ortografisk <sub>2</sub>	Ortografisk <sub>2</sub>

*Note.* Figuren er hentet fra Frith (1985) s. 311

**Figur 1** Friths sekstrinnsmodell for tilegnelsen av lese- og staveferdigheter

Det er først når de logografiske ferdighetene har kommet til trinn 2 at disse kan benyttes i stavingen. De logografiske ferdighetene er altså ”driveren” for utviklingen av staveferdigheter i den første fasen. Den alfabetiske strategien tas først i bruk i stavingen, og blir ”driveren” for videre utvikling av leseferdigheter når den har nådd trinn 2. I fase 3, den ortografiske, er det

igjen lesingen som er ”driveren” for stavingen. De ortografiske ferdighetene må ha nådd trinn 2 for å være presise nok til å kunne nyttiggjøres i stavingen. Hver av de tre fasene består først av et trinn der det er ulikheter i hvilke strategier som benyttes for henholdsvis lesing og staving, og så et trinn hvor den samme strategien benyttes.

Ehri (1995, 2000) beskriver de ulike nivåene i lese- og staveutviklingen på denne måten:

**Føralfabetisk nivå:** På dette nivået vet barna lite om hvordan det alfabetiske systemet representerer muntlig språk. For å huske hvordan de skal lese ord, lagrer de visuelle trekk som for eksempel de to ”bulkene” i m-en i ordet ”kamel”. De kan skrive ord som overfladisk ligner ordentlige ord, men de gjør det ikke ut i fra at bokstavene de skriver ligner på lydene i ordet. De kan på dette nivået ikke lese hva de har skrevet.

**Delvis alfabetisk nivå:** Barna har nå kunnskaper om navn og lyder på bokstaver, og kan lese og skrive ord på et delvis alfabetisk eller semifonetisk nivå. Kunnskapen om det alfabetiske systemet er ufullstendig. Barna kjenner igjen noen av bokstavene i et ord, og gjetter seg til hvilket ord det er ut i fra dette. Barna klarer ikke å avkode ukjente ord, og korrekt staving er vanskelig på dette nivået fordi barna ikke har nok kunnskap om det alfabetiske systemet.

**Fullt alfabetisk nivå:** På dette nivået kan barna dele ord inn i riktige fonemer og de kjenner til sammenhengen mellom grafem og fonem. Denne kunnskapen gjør at de kan dra full nytte av det alfabetiske systemet og avkode ord ganske godt.

**Konsolidert alfabetisk nivå:** Etterhvert som barna får praktisert lesing og konvensjonell staving av ord, lærer de at ord består av større enheter og bokstavkonstellasjoner som man finner igjen i ulike ord. Slike større enheter kan være stavelser, deler av stavelser (som opptakter og rim) og affikser i begynnelsen eller slutten av ord.

Ehri (1995, 2000) påpeker at det å lære seg å lese og stave ord er tett knyttet sammen. Lesing og staving utvikler seg sammen og påvirker hverandre gjensidig; lesingen påvirker stavingen, men på samme tid påvirker stavingen lesingen.

## 2.2 Variabler som kan forklare begynnende lese- og staveutvikling

Verhagen et al. (2008) poengterer at effekten av ulike variabler avhenger av a) hvor konsistent språket er og av b) hvordan både den avhengige variabelen og forklaringsvariablene er målt. Et annet viktig spørsmål er hvorvidt man har kontrollert for andre mulige forklaringsvariabler og hvorvidt man tar høyde for den autoregressive effekten, det vil si effekten av den avhengige variabelen målt på et tidligere tidspunkt (de Jong & van der Leij, 1999). Resultetene vil også avhenge av hvor gamle barna er og hvor i utviklingsløpet de er, samt karakteristikk ved utvalget (dårlige lesere/stavere eller uselekterte barn).

Det er enighet i forskningen om at fonologisk bevissthet er en viktig prediktor for utviklingen av begynnende lese- og staveferdigheter. Også Rapid Automatized Naming (RAN) ser ut til å være en god prediktor for senere leseferdigheter. Når det gjelder staving, er det ikke så mange studier som har sett på betydningen av RAN, og disse studiene gir heller ikke et entydig bilde (Lervåg & Hulme, 2010).

### 2.2.1 Fonologisk bevissthet

Den fonologiske evnen som har fått mest oppmerksomhet i forbindelse med lesing og staving, er fonologisk bevissthet. Fonologisk bevissthet kan defineres som evnen til å oppfatte at muntlige ord består av ulike lydenheter (som fonemer, opptakter, rim og stavelser) og evnen til å kunne manipulere disse lydenhetene (Lervåg, 2005). Fonem er den minste lydenheten i språket som kan forandre et ords mening. Hvis man for eksempel forandrer det første fonemet i "hatt" fra /h/ til /k/, får man ordet "katt". Opptakten i en stavelse er en konsonant eller konsonantkluster som kommer forut for en vokal i stavelsen, for eksempel /br/ i "bror" og /b/ i "bok". Rimet til stavelsen er vokalen og de påfølgende konsonantene, for eksempel /is/ i "gris" og /opp/ i "hopp". En stavelse består av både en opptakt og et rim.

Torgesen, Wagner, Rashotte, Burgess & Hecht (1997) definerer fonologisk bevissthet som bevissthet om, og tilgang til, lydstrukturen i det muntlige språket.

Det er ikke nødvendigvis den fonologiske bevisstheten i seg selv som predikerer utviklingen av begynnende lese- og staveferdigheter. Noen forskere mener at fonologisk bevissthet



reflekterer kvaliteten på de underliggende fonologiske representasjonene. Godt spesifiserte fonologiske representasjoner gjør det enkelt både å manipulere lydenheter (som for eksempel i ulike tester av fonologisk bevissthet) og å lære hvilke bokstaver eller bokstavkonstellasjoner som hører sammen med hvilke lydenheter i talespråket (Lervåg, 2005).

Mange av studiene som har sett på effekten av fonologisk bevissthet på lesing og staving, har ikke vært longitudinelle, og har dermed ikke tatt hensyn til mulige autoregressive effekter (det vil si effekten av lesing og staving målt på et tidligere tidspunkt). Det er spesielt viktig å kontrollere for autoregressive effekter dersom man er interessert i utvikling. Dersom man ikke kontrollerer for lesing og staving på et tidligere tidspunkt, kan man ikke vite om det virkelig er en effekt av fonologisk bevissthet på lesing og staving, eller om denne effekten heller skyldes individuelle forskjeller i lesing og staving på et tidligere tidspunkt (de Jong & van der Leij, 1999).

Flere studier, som også har tatt høyde for de autoregressive effektene, finner at fonologisk bevissthet forklarer utvikling i lesing og/eller staving. I studier av nederlandske barn (de Jong & van der Leij, 1999; Verhagen et al., 2008) og norske barn (Lervåg, Bråten & Hulme, 2009) finner man at fonologisk bevissthet kun har betydning det første året av leseopplæringen. Lervåg & Hulme (2010) finner at en kombinert faktor av fonologisk bevissthet og bokstavkunnskap (betydningen av disse to prediktorene kunne ikke skilles fra hverandre) målt 10 måneder før start av formell lese- og staveopplæring, var gode prediktorer for staveferdigheter over ett år seinere.

Landerl & Wimmer (2008) finner at fonologi målt i begynnelsen av første klasse bidro til å forklare variasjoner i lesing i slutten av første klasse, samt til å forklare variasjoner i staving i fjerde og åttende klasse hos tyske barn. De mener at én forklaring på sammenhengen mellom tidlige fonologiske ferdigheter og senere utvikling av staveferdigheter kan være at det å bygge opp ortografiske representasjoner krever dannelsen av multiple assosiasjoner mellom skrevet og talt ord. Barn med fonologiske vansker er muligens ikke i stand til å etablere slike multiple assosiasjoner.

Leppänen et al. (2006) finner at fonologisk bevissthet predikerer nivået av både lese- og staveferdigheter ved starten av førskolen, og også utviklingen av disse ferdighetene i løpet av dette førskoleåret. Fonologisk bevissthet bidro i denne undersøkelsen ikke til å forklare senere utvikling i lesing eller staving etter at lesing og staving på tidligere tidspunkt var kontrollert

for. Torgesen et al. (1997) fant at de tre variablene fonologisk bevissthet, fonologisk kortidsminne og RAN alle bidro til å forklare vekst i leseferdigheter hos engelskspråklige barn når de var inne som eneste variabel. Når alle tre var inne på en gang, var det kun fonologisk bevissthet som bidro til å forklare vekst i lesing fra andre til femte klasse. Det ser ut til at fonologisk bevissthet har betydning som forklaringsvariabel over en lengre periode i et inkonsistent språk som engelsk enn i mer konsistente språk som norsk, nederlandsk og finsk.

### 2.2.2 Rapid Automated Naming (RAN)

En annen variabel som ser ut til å ha sammenheng med begynnende lese- og staveutvikling, er Rapid Automated Naming (RAN). På slike RAN-tester skal man navngi symboler som bokstaver, tall, farger eller objekter så fort man kan. Naming speed eller rapid serial naming er andre betegnelser som brukes på det samme målet. Det er vanlig å omtale RAN-testene som henholdsvis ikke-alfanumeriske (farger og objekter) og alfanumeriske (tall og bokstaver).

En RAN-test (alfanumerisk) kan for eksempel se slik ut (Lervåg, 2005):

4	2	9	7	6	9	2	6	7	4
6	4	7	2	9	4	6	2	7	9
7	6	2	9	4	6	4	9	7	2
9	6	2	7	4	2	7	9	4	6

Det eksisterer ulike teorier om hva RAN egentlig er et uttrykk for:

- 1) Fonologisk prosessering: RAN-oppgaver måler individuelle forskjeller når det gjelder hvor effektivt visuelle symboler blir omkodet til fonologiske representasjoner. (Logan, Scatschneider, & Wagner, 2011). Det handler altså om hvor raskt og effektivt man har tilgang til leksikon i langtidsmindet.

- 2) Generell kognitiv hurtighet: Teorien om at RAN-oppgaver reflekterer generell kognitiv hurtighet har blitt testet av flere forskere, og de finner at RAN bidrar til å forklare varians i lesing etter at det er kontrollert for generell kognitiv hurtighet (Logan et al., 2011).
- 3) Multiple constructs: RAN-oppgaver tar høyde for en rekke ulike prosesser som tilgang til leksikon, generell kognitiv hurtighet og automatisering, samt visuelle,- artikulatoriske- og oppmerksomhetsprosesser. Den kritiske komponenten som her relaterer RAN til lesing, handler om hvor effektivt alle disse prosessene integreres (Wolf & Bowers, 1999, referert i Logan et al., 2011).
- 4) En underliggende struktur i sentralnervesystemet: Denne underliggende strukturen har betydning for hvor nøyaktig og effektivt man lagrer ortografisk informasjon. Denne strukturen blir sett på som delvis uavhengig av de strukturene som kreves for fonologisk prosessering. RAN kan i dette perspektivet ses på som grunnleggende nevrologiske prosesser som har betydning for både lesing og staving. Lervåg & Hulme (2010) spekulerer i om RAN har sammenheng med nervebaner involvert i identifisering av objekter og benevning, og at dette er de samme nervebanene som fungerer som en kritisk komponent i begynnende lese- og staveutvikling. En manglende effektivitet i denne mekanismen begrenser utviklingen av både lese- og staveferdigheter.

Sammenhengen mellom RAN og lesing ble opprinnelig funnet blant dyslektikere. Studier som viser til en slik sammenheng mellom RAN og lesing også i uselekterte utvalg er blant annet de Jong & van der Leij, 1999; Landerl & Wimmer, 2008; Lervåg & Hulme, 2009 og Verhagen et al., 2008. Disse studiene måler RAN på litt ulike måter. Noen tester både ikke-alfanumerisk og alfanumerisk RAN, mens andre kun bruker én av testene. RAN er i alle disse studiene en god prediktor for leseflyt, også etter det første skoleåret.

I en studie som undersøker vekst i staving og nonordstaving, finner Lervåg & Hulme (2010) at ikke-alfanumerisk RAN målt i starten av første klasse predikerer både staving og nonordstaving over ett år senere.

## 2.3 Variasjoner mellom språk

Det er forskjeller mellom språkene i hvordan lydstrukturen i det talte språket, fonologien, representeres i skriften (ortografien). Dette vil påvirke hvor raskt og effektivt man utvikler lese- og skriveferdigheter, og i hvilke perioder ulike kognitive og lingvistiske variabler predikerer denne utviklingen (Lervåg & Hulme, 2010). Det er viktig å være klar over dette i tolkningen av forskningsresultater, da mye av den eksisterende forskningen innenfor lese- og skrivefeltet er foretatt innenfor en inkonsistent ortografi som engelsk.

Seymour, Aro & Erskine (2003) finner at engelskspråklige barn trenger to og et halvt år og danskspråklige barn to år for å nå det samme nivået på lesenøyaktighet som barn fra konsistente ortografier nådde på ett år. Et eksempel på det sistnevnte er finsk, der både grafem-fonem og fonem-grafem sammenhengen er konsistent. Som nevnt tidligere, er fonem den minste lydenheten i språket som kan forandre et ords mening og grafem er fonemets skriftlige uttrykk. I finsk, der både grafem-fonem og fonem-grafem sammenhengen er konsistent, læres de grunnleggende ferdighetene, som avkoding, relativt raskt. (Parilla et al., 2005). Leppänen, Niemi, Aunola & Nurmi (2004) viser at slike grunnleggende ferdigheter i en konsistent ortografi som finsk er på plass mellom tre og syv måneder etter at man startet med formell lese- og skriveopplæring.

Språk varierer ut i fra dimensjonene konsistent/inkonsistent (shallow/deep) ortografi og enkel/kompleks stavelsesstruktur. Seymour (2005) kategoriserer noen av de europeiske språkene ut i fra disse dimensjonene på denne måten:

1. Enkel stavelsesstruktur og konsistent ortografi (finsk, gresk, italiensk, spansk)
2. Kompleks stavelsesstruktur og konsistent ortografi (tysk, norsk, islandsk, svensk og nederlandsk)
3. Enkel stavelsesstruktur og inkonsistent ortografi (portugisisk, fransk)
4. Kompleks stavelsesstruktur og inkonsistent ortografi (dansk, engelsk)

Norsk er altså kategorisert som en konsistent ortografi, men med en kompleks stavelsesstruktur. Seymour et al. (2003) finner at en slik kompleks stavelsesstruktur påvirker lesing av nonord, men ikke lesing generelt. Hvorvidt en ortografi er konsistent eller ikke, kan

derfor sies å være den dimensjonen som har mest betydning, fordi den påvirker utviklingen når det gjelder lesing av både vanlige ord og nonord.

Det er viktig å skille mellom lesing (forbindelsen mellom staving og lyd) og staving (forbindelsen mellom lyd og staving) når det gjelder graden av konsistens (Lervåg & Hulme, 2010). Graden av konsistens når det gjelder staving (altså mellom fonemer og grafemer) har ikke fått så mye oppmerksomhet i norsk, men det er generelt akseptert at fonem-grafem forbindelsen også er moderat konsistent (Bråten, 1994; Hagtvat & Lyster, 2003, ref. i Lervåg & Hulme, 2010).

## **2.4 Tidligere forskning på forholdet mellom begynnende lese- og staveutvikling**

Mange studier har vist at lese- og staveferdigheter er sterkt relatert til hverandre, men det er ikke klart hvorvidt det er snakk om et resiprokt forhold helt fra begynnelsen av eller om lesing og staving bytter på å påvirke hverandre i ulike faser av utviklingen. Det har heller ikke vært gjennomført så mange longitudinelle studier på dette feltet (Leppänen et al. 2006).

I følge Fitzgerald og Shanahan (2000), er de kunnskapsrepresentasjonene og kognitive ferdighetene som ligger til grunn for lesing og staving veldig like. Det er derfor naturlig å tenke seg at utviklingen av disse to ferdighetene er tett knyttet sammen. Et slikt gjensidig og tett forhold mellom begynnende lese- og staveutvikling har i liten grad vært systematisk underøkt (Leppänen et al., 2006).

Caravolas, Hulme & Snowling (2001) følger engelske barns stave- og leseutvikling de tre første skoleårene. Deres funn bekrefter teorien til Frith (1985) om at det i den alfabetiske fasen er fonologisk staving som forklarer leseutviklingen. Senere i utviklingen finner de at lesing er en sterk prediktor av konvensjonell staving, en ferdighet som krever ortografisk kunnskap utover evnen til fonologisk omkodning. De finner også at staving ikke har innflytelse på lesingen utover den helt tidlige fasen av lese- og staveutviklingen. Også dette er i tråd med Frits hypotese om at i den senere fasen av utviklingen er lesing som forklarer staveutviklingen.

Leppänen et al. (2006), undersøker i en longitudinell undersøkelse hvordan begynnende lese- og staveferdigheter påvirker hverandre i overgangen fra barnehagen til førskolen og så videre til begynnelsen av andre klasse hos finske barn. De undersøker også forløperne til disse ferdighetene. De analyserer sine longitudinelle data ved hjelp av autoregressive modeller med kryssladninger, noe som gjør det mulig å teste retningen på resiproke sammenhenger mellom barns begynnende lese- og staveferdigheter, og på mulige endringer av dette forholdet over tid. De fant at staveferdigheter ved starten av førskolen predikerte leseferdigheter gjennom førskolen og første klasse. Leseferdigheter ved slutten av førskolen og ved slutten av første klasse predikerte staveferdigheter i henholdsvis første og andre klasse. Da fonologisk bevissthet ved starten av førskolen ble kontrollert for, predikerte ikke lenger staveferdigheter ved starten av førskolen leseferdigheter på slutten av dette året. Fonologisk bevissthet bidro altså til å forklare denne tidlige linken mellom lesing og staving.

Disse resultatene reflekterer muligens den finske ortografien. Friths og Ehris teorier ble utviklet med basis i data fra engelskspråklige barn. Leppänen et al. (2006) spør seg om effekten av tidlige staveferdigheter på senere leseferdigheter har mindre betydning i konsistente ortografier så fort barna har oppnådd et visst nivå av fonologisk bevissthet og har startet ”å knekke koden”.

## 2.5 Forskningsspørsmål

Eksisterende forskning tyder på at de kunnskapsrepresentasjoner og kognitive ferdigheter som ligger til grunn for å lære seg å lese og stave er ganske like, og at det derfor er naturlig å tenke seg at utviklingen av disse to ferdighetene er tett knyttet sammen.

Den kognitive variabelen som har fått mest oppmerksomhet når det gjelder lese- og staveutvikling, er fonologisk bevissthet. Fonologisk bevissthet ser ut til å størst betydning tidlig i utviklingen (Lervåg, Bråten & Hulme, 2009; de Jong & van der Leij, 1999, Leppänen et al., 2006). Betydningen av fonologisk bevissthet for staving ser ut til å gjelde for en noe lengre tidsperiode enn for lesing (Landerl & Wimmer, 2008). Rapid Automated Naming ser ut til å ha betydning for leseflyt også utover det helt tidlige utviklingsstadiet (Verhagen et al., 2008; Lervåg & Hulme, 2009). Når det gjelder betydningen av RAN for utvikling av staveferdigheter, er ikke bildet fullt så tydelig.

Selv om det er enighet om at utviklingen av lesing og staving er tett knyttet sammen, finnes det ulike teorier om hvordan dette forholdet er. I følge Frith (1985) skal fonologisk staving forklare fonologisk lesing først, og ortografisk lesing forklare ordstaving senere i utviklingen. I følge Ehri (1995, 2000) skal det være en mer generell, gjensidig påvirkning mellom lesing og staving hele veien.

I datamaterialet som ligger til grunn for analysene i denne oppgaven, er lesing og staving målt gjennom lesing og staving av enkeltord, både nonord og konvensjonelle ord. Lesing og staving av nonord regnes for å være et godt mål på tidlige ordavkodingsferdigheter, og kan sies å være mer ”fonologiske i sin natur” enn ortografisk lesing og konvensjonell ordstaving.

Med bakgrunn i dette stilles følgende forskningsspørsmål:

- 1) Er det nonordstaving som forklarer nonordlesing tidlig i utviklingen og ortografisk lesing som forklarer ordstaving litt senere i utviklingen, eller er det slik at lesing og staving utvikler seg sammen og påvirker hverandre gjensidig?

Disse hypotesene testes ut både med og uten fonologisk bevissthet og alfanumerisk RAN som kontrollvariabler (kovariater). Videre spør jeg om

- 2) hvor godt autoregressive modeller og latente vekstkurvemodeller egner seg til å si noe om den gjensidige utviklingen av begynnede lese- og staveferdigheter og forholdet mellom disse.

## 3 Metode

Mange teoretiske hypoteser fokuserer på forandring, og man må ha longitudinelle data for å fullt ut evaluere slike hypoteser og for å kunne si noe om retningen når det gjelder mulige årsaksforhold. Det eksisterer en rekke ulike teknikker for analyse av longitudinelle data. Hvilken tilnærming man velger, avhenger av teorien som ligger til grunn og hvordan forutsetningene for bruk av de ulike analysemetodene samsvarer med de empiriske dataene (Bollen & Curran, 2006). Jeg har valgt to metoder som er vanlig å bruke i analyse av longitudinelle data, nemlig autoregressive modeller og latente vekstkurvmodeller.

Fordi både latente vekstkurvmodeller og autoregressive modeller er varianter av Structural Equation Modeling, SEM, vil jeg redegjøre for temaer som latente variabler, modelltilpasning, evaluering av modeller, estimering av parametre i modellen og manglende data. Med dette som bakgrunn, vil jeg så redegjøre mer spesifikt for latente vekstkurvmodeller og autoregressive modeller.

### 3.1 Strukturelle likningsmodeller

Strukturelle likningsmodeller (SEM) brukes som en fellesbetegnelse for et generelt analysesystem der hensikten er å undersøke relasjoner mellom målte variabler og/eller latente variabler, basert på et sett av observerte indikatorvariabler (Hagtvet, 2010).

SEM-modeller består av to komponenter: målemodellen og den strukturelle modellen. Målemodellen spesifiserer antall faktorer, hvordan de ulike indikatorene er relatert til de latente faktorene, og forholdet mellom indikatorfeilene. Den strukturelle modellen spesifiserer hvordan de latente faktorene er relatert til hverandre (Brown, 2006). Det er vanlig å estimere målingsmodellen først. “Dersom målingsmodellen ikke representerer de latente variablene tilfredsstillende, kan det føre til feilestimering av regresjonsparametrene i strukturmodellen” (Hagtvet, 2010, s. 287).

Som nevnt over, er både latente vekstkurvmodeller og autoregressive modeller applikasjoner av Structural Equation Modeling, SEM, og de har dermed de samme fordelene. Noen av disse fordelene er muligheten til å benytte maximum likelihood teknikker for manglende data, å



kunne vurdere modelltilpasningen og å kunne identifisere mulige kilder til dårlig tilpasning, samt muligheten til å inkludere latente kovariater og latente repeterte variabler i modellene (Bollen & Curran, 2006).

I vanlige SEM-analyser er det kovariansene som analyseres. Det innebærer at alle observerte variabler er sentrerte, og de latente variablene har et gjennomsnitt på 0. Dette er i noen analysesammenhenger en stor mangel. Gjennomsnittet til repeterte målinger forventes noen ganger å endre seg over tid. Gjennomsnitt kan i SEM estimeres ved å legge til en gjennomsnittsstruktur i tillegg til modellens kovariansstruktur. Input-data for analyse av modeller med en gjennomsnittsstruktur er kovarianser (eller korrelasjoner og standardavvik) og gjennomsnitt (Kline, 2005).

### **3.1.1 Latente variabler**

En av de store fordelene med SEM er at man kan estimere forholdet mellom variabler justert for målefeil ved å benytte seg av latente variabler. En latent variabel er en variabel som ikke kan observeres direkte.

Målingsmodellen som jeg refererte til i avsnittet over, bygger på en teori om hvordan skårer fra observerte indikatorer som representerer et begrep, er sammensatt. Ved hjelp av en konfirmerende faktoranalyse, kan man finne den komponenten som alle indikatorene har felles. ”Man kan anta at felleskomponenten er en variabel som reflekterer den latente variabelen man ønsker å dra slutninger om. Faktoranalyse er derfor en latent variabelanalyse som estimerer innslaget av felleskomponenten og feilkomponenten i de observerte indikatorene” (Hagtvet, 2010, s. 288).

En analysemetode som for eksempel multippel regresjon, forutsetter at variablene er målt uten feil og at all varians i et observert mål er sann varians. Denne forutsetningen hoder sjelden i samfunnsvitenskapen. Når man benytter latente variabler, blir den variansen indikatorene har felles, operasjonalisert som sann skåre varians, som blir “overført til” den latente variabelen (Brown, 2006). Dette medfører at estimerer av sammenhengen mellom begreper man ønsker å få vite noe om, blir riktigere.

### 3.1.2 Estimering av parametre i modellen

Begrepet modellering står sentralt i SEM-metodologien. Til forskjell fra andre analyseformer er SEM-analyser ikke primært dataorienterte, men konseptuelt orienterte. Man lager modeller ut i fra begreper og teorier som inngår i forskningsproblemet (Hagtvet, 2010). Dette er den teoretiske modellen, " $\Sigma$ ". Målet er å få estimater for hvert enkelt parameter i den teoretiske modellen " $\Sigma$ " som gjenskaper varians-kovariansmatrisen (symbolisert med " $S$ ") i utvalget best mulig.

Estimeringen er en matematisk operasjon der målet er å minimalisere forskjellen mellom  $\Sigma$  og  $S$  (fitting function). Den mest brukte estimeringsmetoden er maximum likelihood (heretter omtalt som ML).

Det underliggende prinsippet i ML-estimering er å finne modellparametre som maksimerer sannsynligheten for å observere de tilgjengelige dataene dersom dataene ble samlet inn fra den samme populasjonen igjen. Dette er en iterativ prosedyre. Dataprogrammet man bruker, som for eksempel Mplus, begynner med et sett av parameterestimater, også kalt startverdier. Disse startverdiene kan genereres automatisk av det dataprogrammet man bruker, eller spesifiseres av forskeren selv. Gjennom den iterative prosedyren redefineres verdien på de ulike parameterstimatene i den hensikt å minske differansen mellom  $\Sigma$  og  $S$ . Hver slik redefinering kalles en iterasjon. Modellen konvergerer når dataprogrammet har kommet frem til et sett av parameterestimater der ytterligere iterasjoner ikke fører til at differansen mellom  $\Sigma$  og  $S$  blir mindre (Brown, 2006). Hagtvet (2010, s. 295) omtaler dette som "å redusere diskrepansen mellom  $\Sigma$  og  $S$  inntil tilstrekkelig likhet er oppnådd. Tilstrekkelig likhet innebærer at forandringen i den anvendte fit-funksjonen,  $F(S, \Sigma)_{ML}$ , er blitt svært liten (f.eks 0.000001)."

ML-estimering forutsetter 1) at utvalgsstørrelsen er stor, 2) at indikatorene er målt på en kontinuerlig skala og 3) at fordelingen til indikatorene er multivariat normal. Brudd på kravet om normalfordeling kan føre til gale standardfeil og dermed påvirke signifikanstesting. (Brown, 2006). Når dataene ikke er normalfordelte, bør man bruke en mer robust estimator enn ML. To slike robuste estimators er MLM og MLR. MLM forutsetter komplette datasett og fungerer dermed ikke dersom man har manglende data. Det gjør derimot MLR (Muthén & Muthén, 1998-2010).

### 3.1.3 Modelltilpasning

Modelltilpasningen sier ikke noe om de enkelte komponenter i modellen, men er et overordnet mål for hvor godt den teoretiske modellen samsvarer med data. De ulike målene for overordnet modelltilpasning er i hovedsak basert på hvor godt den teoretiske modellen reproduserer de observerte variabelens varianser, kovarianser og gjennomsnitt (Bollen & Curran, 2006).

Overidentifiserte modeller, det vil si modeller med flere observasjoner enn estimerte parametre (og dermed et positivt antall frihetsgrader), passer normalt ikke perfekt til dataene. Dermed oppstår behovet for å vurdere hvor godt eller dårlig en modell beskriver dataene. Det er viktig å ikke la seg blinde av en god modelltilpasning. Man må også vurdere de ulike parameterestimaterne; en god modelltilpasning betyr ikke nødvendigvis at modellen er teoretisk fornuftig. Det er heller ikke sikkert at en modell med god tilpasning forklarer så mye av det vi er ute etter å vite noe om. Uforklart varians hos en eller flere endogene variabler kan altså være stor selv om modelltilpasningen ser ut til å være god (Kline, 2005).

Det finnes flere ulike indekser for modelltilpasning, og det er viktig å være klar over at de måler ulike aspekter ved modelltilpasningen. Det er derfor vanlig å se på og rapportere flere forskjellige indekser. Jeg vil i denne oppgaven rapportere  $\chi^2$  (kji-kvadrat) med Yuan-Bentler korreksjoner på grunn av non-normale data, Root Mean Square Error of Approximation (RMSEA) med 90% konfidensintervall, Comparative Fit Index (CFI) og Standardized Root Mean Square Residual (SRMR).

## Kji-kvadrat

Kji-kvadrat ( $\chi^2$ ) er en teststatistikk der den oppnådde minimumsverdien av den tidligere omtalte fit-funksjonen transformeres til en kji-kvadratverdi ved hjelp av følgende formel:

$\chi^2_{df} = (N - I) F_{ML}$ , der  $df$  er antall frihetsgrader,  $N$  er utvalgsstørrelsen og  $F_{ML}$  er minimumsverdien av maximum likelihood fit-funksjonsverdien.

For å kunne benytte denne teststatistikken til å vurdere modelltilpasningen, må vi ha en overidentifisert modell (en modell med positivt antall frihetsgrader). I en slik situasjon betyr et statistisk signifikant resultat at den teoretiske modellen ikke eksakt reproducerer gjennomsnittene eller kovariansmatrisen til de observerte variablene (Bollen & Curran, 2006). “Modelltilpasningen vurderes som god dersom kji-kvadratverdien ikke er signifikant forskjellig fra null, som impliserer at modell uttrykt ved  $\Sigma$  og data representert ved  $S$  ikke er systematisk forskjellige.” (Hagtvet, 2010). Det er flere grunner til at man ikke bare skal basere seg på  $\chi^2$  teststatistikk i vurderingen av overordnet modelltilpasning: Dersom de observerte variablene ikke er normalfordelte, kan teststatistikken bli for høy eller for lav. Det finnes korrigeringer som tar høyde for dette. En annen svakhet ved dette tilpasningsmålet, er at den statistiske styrken er stor i store utvalg. Statistisk styrke sier noe om sannsynligheten for å forkaste en falsk nullhypotese. I store utvalg kan testresultatet bli statistisk signifikant selv om forskjellen mellom den teoretiske modellen og den estimerte modellen ikke er stor. Store utvalg har altså statistisk styrke til å avdekke selv små feil i modellspesifiseringen (Bollen & Curran, 2006).

Det er verdt å merke seg at  $\chi^2$  er nyttig i sammenligningen av hierarkiske modeller. For at man skal kunne bruke  $\chi^2$  til å sammenligne modeller, må modellene som sammenlignes være nøstet i hverandre. Nøstet betyr at parametrene til den nøstede modellen er en restriktiv form av parametrene til sammenligningsmodellen. Et eksempel på dette er to modeller som er identiske med unntak av at noen parametre for eksempel er satt til 0 eller satt til å være like, mens de i den andre modellen er frie parametre. Dersom man skal sammenligne to modeller som ikke er nestet i hverandre, kan man benytte seg av tilpasningsmål som Akaike information criterion, AIC, eller Bayesian information criterion, BIC. For begge disse målene indikerer mindre verdier bedre tilpasning enn høye verdier (Bollen & Curran, 2006).

Det finnes andre tilpasningsmål enn  $\chi^2$  som ikke er så sensitive for utvalgsstørrelse.

## Root Mean Square Error of Approximation

RMSEA er en såkalt "stand-alone" indeks. Det vil si at det ikke foretas en sammenlikning med en baseline-modell (Bollen & Curran, 2006). RMSEA favoriserer den enkleste modellen dersom man har to ulike modeller med samme overordnede forklaringskraft. Desto nærmere 0 RMSEA er, desto bedre modelltilpasning. En tommelfingerregel er at  $RMSEA < .05$  indikerer god tilpasning, verdier mellom  $.05$  og  $.08$  indikerer grei tilpasning og  $RMSEA > .10$  indikerer dårlig tilpasning. (Browne & Cudeck, 1993, referert i Kline, 2005). En av fordelene med RMSEA er at den oppgis sammen med 90% konfidensintervall (Bollen & Curran, 2006).

## Comparative Fit Index

CFI er en såkalt "baseline fit" indeks. Denne typen tilpasningsmål vurderer den relative forbedringen i modelltilpasningen for forskerens modell sammenliknet med en baselinemodell, det vil si en modell der man antar ingen relasjoner mellom variablene. En tommelfingerregel er at dersom denne typen indekser har en verdi større enn  $.95$ , indikerer det en grei modelltilpasning. Desto nærmere verdien er  $1.0$ , desto bedre. Men en verdi på  $1.0$  indikerer ikke at modellen har perfekt fit, men at  $\chi^2$  er mindre enn antall frihetsgrader (Kline, 2005).

## Standardized Root Mean Square Residual

Et tilpasningsmål, som i motsetning til RMSEA og CFI ikke er relatert til  $\chi^2$ -kvadratverdien, er SRMR. Denne indeksen er en oppsummering av hvor store residualene er i kovariansstrukturen, nærmere bestemt den gjennomsnittlige standardiserte differansen mellom den observerte ( $S$ ) og den teoretiske ( $\Sigma$ ) kovariansmatrisen. Jo nærmere 0 denne verdien er, desto bedre (Brown, 2006).

## Kombinasjon av tilpasningsmål

Hu & Bentler (1999) foreslår å kombinere minst to ulike tilpasningsmål; SRMR i kombinasjon med enten RMSEA eller CFI. Også andre kombinasjoner er mulig. Slike kombinasjoner ser ut til å beskytte bra mot både "type 1"-feil (det vil si å forkaste en korrekt nullhypotese) og "type 2"-feil (det vil si å beholde en gal nullhypotese).

Hu & Bentler (1999) foreslår følgende kombinasjoner av de tilpasningsmålene jeg har valgt å rapportere i mine analyser:

SRMR < .09 i kombinasjon med CFI > .95-.96

eller

SRMR < .06 i kombinasjon med RMSEA < .06.

Den siste kombinasjonen blir ikke anbefalt å bruke dersom utvalgsstørrelsen er under 250 fordi RMSEA ser ut til å forkaste en sann modell for ofte under slike betingelser. SRMR anbefales fordi den ser ut til å fungere bra i forhold til feilspesifiserte kovarianser, og RMSEA og CFI fordi de ser ut til å fungere bra i forhold til feilspesifiserte faktorladninger.

### 3.1.4 Evaluering av modeller

Når det gjelder evaluering av modeller, er en av de viktigste aspektene det som skjer i forkant av de statistiske analysene. Man må argumentere for at modellen er meningsfull og brukbar med bakgrunn i tidligere forskning og teori (Brown, 2006). Etter at modellen er rettfærdiggjort substansielt, bør modellen vurderes ut i fra følgende tre aspekter:

- 1) Overordnet modelltilpasning (er redegjort for i avsnittet over).
- 2) Om det er enkelte deler av modellen som har dårlig fit.
- 3) Parameterestimatenes styrke, signifikans og retning.

Den overordnede modelltilpasningen er en global deskriptiv oppsummering av modellens evne til å reprodusere kovariansmatrisen fra rådataene. De to andre aspektene bidrar med mer spesifikk informasjon om modellen og hvor brukbar og akseptabel den er, vurdert opp mot tidligere forskning og teori (Brown, 2006).

To statistikker som kan brukes til å identifisere eventuelle deler av modellen med dårlig tilpasning er modifikasjonsindeksen og residualmatrisen:

*Modifikasjonsindeksen* er et anslag av hvor mye den overordnede  $\chi^2$ -verdien vil reduseres dersom man estimerer et parameter.

*Residualmatrisen* gir spesifikk informasjon om hvor bra hver eneste varians og kovarians har blitt reprodusert av parameterestimaterne i modellen. Store positive (standardiserte) residualer betyr at kovariansen underestimeres og kan indikere at det trengs flere parametre i modellen for å ta høyde for kovariansen mellom indikatorene. Store negative residualer innebærer at kovariansen mellom to indikatorer overestimeres. “Stor” i denne sammenhengen er en verdi på den standardiserte residualen på 1.96 (positiv eller negativ). Som for  $\chi^2$ -teststatistikken må man her være klar over at utvalgsstørrelsen spiller inn. I store utvalg kan man få ”for store” standardiserte residualer selv om avvikene egentlig ikke er så store.

Det er viktig å være klar over at både residualer og modifikasjonsindeks er sensitive for utvalgsstørrelse. Nok en gang er det viktig å poengtere at det å løse opp et parameter (altså å estimere det) må begrunnes substansielt:

Especially when sample size is large, modification indices and standardized residuals will often indicate the presence of parameters that, if freed, would improve the fit of the model.

However, such parameters should not be freed with the sole intent of improving model fit.

Rather, all model respecifications must be justified on the basis of prior research and theory (Brown, 2006, s. 122-123).

### 3.1.5 Manglende data

Det er meget sjelden man som forsker har datasett der man har data på alle variabler for alle case. De strategiene som ofte brukes for å håndtere manglende data, som listwise eller pairwise deletion, er ikke hensiktsmessige: de fører til tap av statistisk styrke og bias i parameterestimer, standardfeil og teststatistikk (Brown, 2006).

Data kan mangle på ulike måter. De kan være “missing completely at random”, MCAR. Dette innebærer at sannsynligheten for manglende data på Y ikke henger sammen med Y eller med verdier på noen av de andre variablene i datasettet. Dette er ofte en litt for streng antakelse. En antakelse som ikke er like streng, er at data er “missing at random”, MAR. Det vil si at manglende data på Y avhenger av verdien på X, men ikke er relatert til verdien på Y når X holdes konstant. Selv om dette er en svakere forutsetning enn MCAR, er det mer sannsynlig at denne forutsetningen holder i et datasett. Dersom data oppfyller forutsetningene for MCAR eller MAR, trenger man ikke ta hensyn til de dataene som mangler (ignorable missingness).

Manglende data kan derimot ikke ignoreres dersom case med manglende data på en gitt variabel har høyere eller lavere verdier på denne variabelen enn de casene som har data på denne variabelen, kontrollert for alle andre variabler i datasettet (Brown, 2006).

Som nevnt tidligere, er strategier for å håndtere manglende data som listwise eller pairwise deletion, ikke hensiktsmessige. Det er nå utviklet mer sofistikerte metoder for å håndtere manglende data, som for eksempel direct Maximum likelihood (ML) og Multiple Imputation (MI). "Perhaps most valuable in these methods is that they are valid under the less restrictive assumptions of MAR rather than MCAR" (Bollen & Curran, 2006, s. 72).

Direkte Maximum likelihood (ML) benytter seg av alle tilgjengelige data, det vil si at N er total utvalgsstørrelse, inkludert case med manglende data. Når direkte ML blir brukt i en manglende data kontekst, blir den ofte referert til som full information maximum likelihood, FIML. Direkte ML forutsetter at data er MCAR eller MAR og multivariat normalfordelte. Full information maximum likelihood, FIML, er default i Mplus. Denne fungerer både når data er normalfordelte (ML) og ikke normalfordelte (MLR) (Muthen & Muthen, 1998-2010).

I de fleste tilfeller representerer direkte ML den beste og enkleste måten å håndtere manglende data på i SEM-analyser, men Multiple Imputation (MI) er et godt alternativ til direkte ML når man ikke har tilgang til analyseprogram som kan bruke direkte ML. Multiple Imputation vil si at manglende elementer i datamatriksen blir fylt inn (imputed) ved hjelp av en metode som tilfeldig velger verdier fra skårer til liknende case. Denne prosessen gjentas flere ganger, og det beregnes så et gjennomsnitt. Dersom denne prosessen kun gjøres en gang, vil standardfeilene bli for små fordi beregningen av standardfeilene ikke vil ta høyde for den ekstra variasjonen i de imputerte dataene. Dette løses ved å gjenta imputasjonsprosessen flere ganger (Brown, 2006).



## 3.2 Latente vekstkurvemodeller

A fundamental observation in the social and behavioral sciences is that people change over time, but not necessarily in the same way or at the same rate. There are individual differences in the rate and direction of change in many contexts, and these individual differences in change are often of scientific and practical interest (Preacher et al., 2008, s. 1).

I latente vekstkurvemodeller, blir det tatt høyde for både faktorgjennomsnitt og varianser. Denne kombinasjonen av analyse både på gruppe- og individnivå er unik for denne analysemetoden (Duncan, T. E, Duncan, S.C & Strycker, 2006).

Latent er noe som ikke kan observeres direkte. En latent vekstkurve observeres indirekte ved bruk av repeterte målinger, og denne vekstenkurven kan være forskjellig fra case til case: Hva er den underliggende veksten som førte til de repeterte målingene for hvert enkelt barn? Fokuset er ikke på leseferdigheter på ett tidspunkt som avhengig av leseferdigheter på et annet tidspunkt for alle barn, men på estimering av en underliggende vekstkurve for alle måletidspunkter for hvert enkelt barn. Målet blir da å estimere en linje, en vekstkurve, som passer best mulig til de repeterte observasjonene over tid for hvert enkelt barn (Bollen & Curran, 2006).

Den underliggende ideen for latente vekstkurvemodeller er at individuell vekst (i for eksempel leseferdigheter) er en funksjon av et latent intercept, en latent slope og tilfeldige feil. Den latente interceptfaktoren sier noe om den gjennomsnittlige startverdien i den longitudinelle utviklingsprosessen og den latente slopefaktoren reflekterer den gjennomsnittlige individuelle veksten over tid. Variansen til intercept- og slopefaktorene sier noe om individuelle variasjoner rundt dette gjennomsnittet (Bollen & Curran, 2006; Preacher et al., 2008; Schlueter, Davidov & Schmidt, 2007).

Latente vekstkurvemodeller gjør det mulig å modellere intraindividuell forandring (i personen) og interindividuell forandring (mellom personer) over tid, og det er også mulig å undersøke mulige forklaringer av denne forandringen (Preacher et al., 2008).

Viktige spørsmål innenfor denne analysemetoden er blant annet:

- Hva er “formen” til den gjennomsnittlige trenden over tid?
- Predikerer startnivå hvordan man vokser?

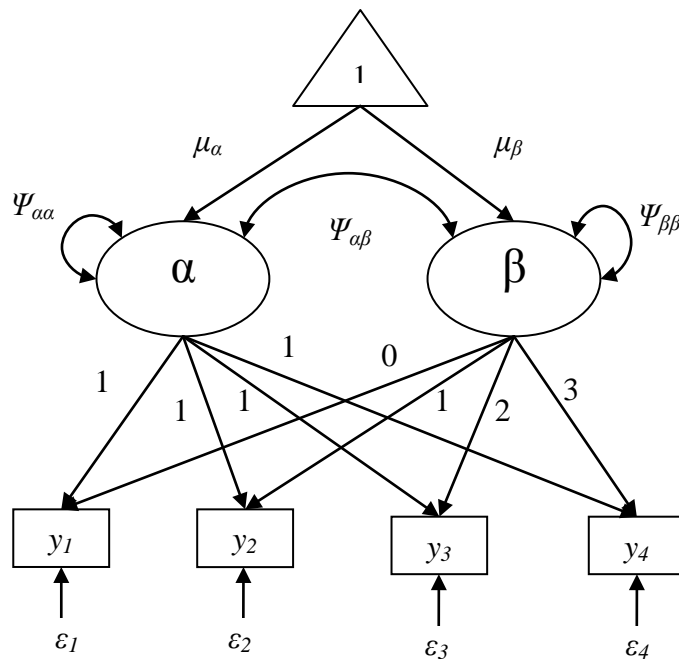
- Hvilke variabler er assosiert med endring over tid?
- Eksisterer det signifikant variasjon i formen på vekstkurven mellom personer?

### 3.2.1 Lineær latent vekstkurvmodell

I en lineær latent vekstkurvmodell er det to underliggende latente faktorer. Disse latente faktorene blir noen ganger referert til som vekstfaktorer. Den første latente faktoren representerer interceptkomponenten til vekstkurven og den andre latente faktoren representerer slopekomponenten til vekstkurven. De observerte repeterte målingene er relatert til de underliggende latente faktorene gjennom faktorladninger.

En modell uten forklaringsvariabler kalles en ubetinget modell. Ved å estimere en slik modell, kan man undersøke om det er signifikant varians i intercept- og slopefaktorene. Signifikant varians i begge disse faktorene betyr at det er variasjoner mellom individer både i hvor de starter og hvordan de vokser. Finner man slike signifikante variasjoner mellom individer, er det interessant å forsøke å finne ut hva som kan forklare disse variasjonene ved å trekke inn forklaringsvariabler. I det vi har slike mulige forklaringsvariabler med i modellen, har vi en betinget latent vekstkurvmodell.

En ubetinget latent vekstkurvmodell illustreres i figur 2.



Figuren er delvis etter Bollen & Curran (2006) og delvis etter Preacher et al. (2008)

**Figur 2** Ubetinget, lineær latent vekstkurvmodell for fire måletidspunkter

En nærmere omtale av de ulike elementene i figur 2 gis nedenfor, og er hentet fra Bollen & Curran (2006).

En ubetinget lineær vekstkurvmodell kan uttrykkes matematisk med likningen

$$y_{it} = \alpha_i + \lambda_t \beta_i + \varepsilon_{yit}$$

$y_{it}$  er vekstkurvevariabel for person  $i$  på tidspunkt  $t$ ;  $\alpha_i$  er random intercept til vekstkurven for person  $i$ . Intercept kan videre uttrykkes med likningen  $\alpha_i = \mu_\alpha + \zeta_{\alpha i}$ , der  $\mu_\alpha$  er gjennomsnittet til interceptene og  $\zeta_{\alpha i}$  er residualen; det vil si avvik fra gjennomsnittlig intercept for person  $i$ .  $\lambda_t$  er et parameter som representerer tid: I en lineær vekstkurvmodell med lik avstand mellom alle måletidspunktene vil  $\lambda_t$  være lik  $t - 1$  for alle  $t$ .  $\beta_i$  er random slope til vekstkurven for person  $i$ . Slope kan uttrykkes med likningen  $\beta_i = \mu_\beta + \zeta_{\beta i}$ , der  $\mu_\beta$  er gjennomsnittet til slopene, og  $\zeta_{\beta i}$  er residualen; det vil si avvik fra gjennomsnittlig slope for person  $i$ .  $\varepsilon_{yit}$  er residualen på hvert måletidspunkt  $t$  for hvert individ  $i$ . Gjennomsnittene  $\mu_\alpha$  og

$\mu_\beta$  estimeres ved å regressere henholdsvis den latente interceptfaktoren og den latente slopefaktoren på en konstant, symbolisert ved trekanten øverst i figur 2.

Kombinert får vi likningen  $y_{it} = (\mu_\alpha + \lambda_t \mu_\beta) + (\zeta_{ai} + \lambda_t \zeta_{\beta i} + \varepsilon_{yit})$ , der første del kalles fixed components og representerer gjennomsnittsstrukturen (mean structure) mens andre del kalles random components og representerer ulike kilder til individuell variasjon.

Videre har vi  $\Psi_{\alpha\alpha} = \text{VAR}(\zeta_{ai})$  som er variansen til interceptene. En signifikant varians betyr at det er variasjoner mellom individer ved starttidspunktet for målingene.  $\Psi_{\beta\beta} = \text{VAR}(\zeta_{\beta i})$  er variansen til slopene. En signifikant varians betyr at det er variasjoner mellom individer i hvordan de vokser.  $\Psi_{\alpha\beta} = \text{COV}(\zeta_{ai}, \zeta_{\beta i})$  er kovariansen til interceptene og slopene og  $\theta_{\varepsilon_{it}} = \text{VAR}(\varepsilon_{it})$  er variansen til residualen på tidspunkt t for individ i. Estimatet av residualvariansen til de gjentatte målingene er variansen i et gjentatt mål som ikke forklares av den underliggende vekstprosessen (Bollen & Curran, 2006).

Det er ulike måter å kode tid på, og man bør kode tid på en måte som gjør det enklest mulig å besvare substansielle spørsmål (Preacher et al., 2008). Alle faktorladningene fra intercept til de gjentatte målingene er i figur 2 satt til 1. Dette reflekterer at hvert enkelt individs intercept er konstant over de gjentatte målingene. Faktorladningen fra slope til de gjentatte målingene gjenspeiler en lineær vekst med likt intervall mellom alle målingene, og der det første måletidspunktet er satt til 0. Dette reflekterer at intercept her er nivået av for eksempel leseferdigheter ved første måletidspunkt. Det er som sagt ulike måter å kode tid på, avhengig av forskningsspørsmålet. Preacher et al. (2008) eksemplifiserer dette ved å vise til et eksempel (utskrivelse fra et rehabiliteringsprogram for rusavhengighet) der det er det siste måletidspunktet som er av størst interesse. Da kodes det siste måletidspunktet som 0, og de foregående fire målingene som henholdsvis -4, -3, -2, og -1. Faktorladningene fra slope øker lineært, og intercept blir definert til å ligge til det femte og siste måletidspunktet.

I mine analyser er det et halvt år mellom T2-T3 og T3-T4 og ett år mellom T4 og T5. T2 er kodet til 0, og er dermed intercept i modellen. Intercept reflekterer gjennomsnittet på T2 (det første måletidspunktet). Siden det bare er et halvt år mellom T2 og T3, og mellom T3 og T4, er de kodet som henholdsvis .5 og 1. Det er ett år mellom T4 og T5, og T5 er derfor kodet som 2.

It is often a good idea to not only collect more than five repeated measures but also to space measurements more closely together during periods when change is occurring most rapidly. This helps avoid estimation problems and to more accurately estimate parameters characterizing change (Preacher et al., s. 13).

Det er viktig å være klar over at forut for estimering av vekstkurven, må man være sikker på at modellen kan identifiseres. Identifikasjon handler om hvorvidt det er unike verdier for estimering av alle parametrene i modellen. Antallet parametre som skal estimeres, må altså være mindre enn antallet observasjoner (Kline, 2005). For at en lineær vekstkurve skal kunne identifiseres, må man ha minimum tre repeterte målinger (Bollen & Curran, 2006).

### **3.2.2 Ikke-lineære latente vekstkurvmodeller**

“Choice of an appropriate shaped trajectory to represent true individual change is an important first step in any analysis of change” (Willett, 1997).

En av de store fordelene med latente vekstkurvmodeller, er at man kan modellere ikke-lineær vekst. Man kan for eksempel ha en hypotese om at grunnleggende leseferdigheter utvikler seg raskt i starten, for så å avta etterhvert. Eller man kan ha en hypotese om at den variabelen man studerer reduseres ganske sakte til å begynne med, for så å reduseres kraftigere etterhvert som tiden går. Uavhengig av hvilken form vekstkurven har, er det mange situasjoner der en modell med lineær forandring ikke samsvarer med den teoretiske modellen man har som utgangspunkt (Bollen & Curran, 2006).

Det er flere ulike måter å modellere en slik ikke-lineær vekst på. Man kan for eksempel ha med en kvadratisk faktor i modellen. I slike modeller representerer slopefaktoren den lineære veksten, mens den kvadratiske faktoren representerer hvordan veksten etterhvert enten vokser fortere eller avtar.

Man kan også modellere ikke-lineær vekst ved å estimere en eller flere av faktorladningene fra slope fritt. En slik modell omtales blant annet som ”completely latent”. Tolkningen av slike modeller varierer ut i fra hvilke faktorladninger som estimeres fritt.

Med en kvadratisk faktor inne i modellen, trenger man fire måletidspunkter for at modellen skal kunne identifiseres (Bollen & Curran, 2006).

Det kan noen ganger være vanskeligere å tolke parametrene i en kvadratisk vekstkurvmodell. Tolkningen av interceptfaktoren er den samme som i en lineær modell. Tolkningen av slopefaktoren er derimot annerledes:

The linear component of the quadratic model described the instantaneous rate of change at the initial assessment. In other words, the linear component of the trajectory function is equal to the slope of the tangent line of the curve, where time is equal to zero. A 1-unit increase in time will lead to a change in the slope of this tangent line: this rate of change per unit is reflected in the quadratic component of the trajectory function. Thus larger (absolute) values of the quadratic component reflect more rapid change in the curvature per unit change of time (Bollen & Curran, 2006, s. 94).

### **3.2.3 Bivariat latent vekstkurvmodell**

Utgangspunktet for en bivariat latent vekstkurvmodell er en hypotese om at to sett av repeterte målingene er relatert over tid. Men denne relasjonen gjelder bare random vekstkomponenter, altså variansen i vekstfaktorene. Det vil si at en eneste direkte sammenhengen mellom de to variablene over tid, er kovariansstrukturen mellom random intercept og random slope. En ubetinget multivariat latent vekstkurvmodell tillater direkte testing av relasjonen mellom to vekstprosesser på random coefficients nivå. Når man har med en eller flere prediktorvariabler i en slik modell, får hver latent vekstfaktor en residual. Residualen er den delen av variansen til hver av de latente faktorene som ikke forklares av prediktorvariabelen eller prediktorvariablene. Kovariansstrukturen mellom de latente faktorene er på nivå av denne residualvariansen. Kovariansen reflekterer nå den lineære sammenhengen mellom den uforklarte variansen for hver vekstfaktor (Bollen & Curran, 2006).

I tillegg til eksogene prediktorer, kan man ha regresjoner mellom de ulike faktorene. Istedenfor å estimere kovarians mellom for eksempel interceptfaktoren for ortografisk lesing og slopefaktoren for staving, kan man estimere en regresjon. Interceptfaktorene blir nå mediatorvariabler for de eksogene variablene i prediksjonen av de latente slopefaktorene for hvert repeterte mål (Bollen & Curran, 2006).

I en bivariat latent vekstkurvemodell er det ikke estimert noen direkte relasjoner mellom mellom de tidsspesifikke målene innen eller mellom de to ulike seriene med repeterte målinger. En slik modell tillater altså ikke tidsspesifikk påvirkning fra for eksempel ortografisk lesing til ordstaving eller omvendt. De relasjonene man antar eksisterer mellom de ulike repeterte målene, blir estimert kun på vekstkurvenivå. Det vil si at veksten generaliseres over alle måletidspunkter (Bollen & Curran, 2006).

### 3.3 Autoregressive modeller

Selv om latente vekstkurvemodeller er en fleksibel tilnærming for å analysere endring over tid, finnes det også andre SEM-baserte modeller som kan brukes til denne typen analyser. En alternativ modell er den autoregressive modellen (Markov simplex modell) der hvert repeterte mål blir sett på som en funksjon av det foregående målet, samt en tidsspesifikk residual. En slik autoregressiv modell tar ikke for seg utviklingen av gjennomsnittsstrukturen over tid, men derimot forklaring av varians på hvert måletidspunkt ved hjelp av foregående måletidspunkt (Preacher et al., 2008).

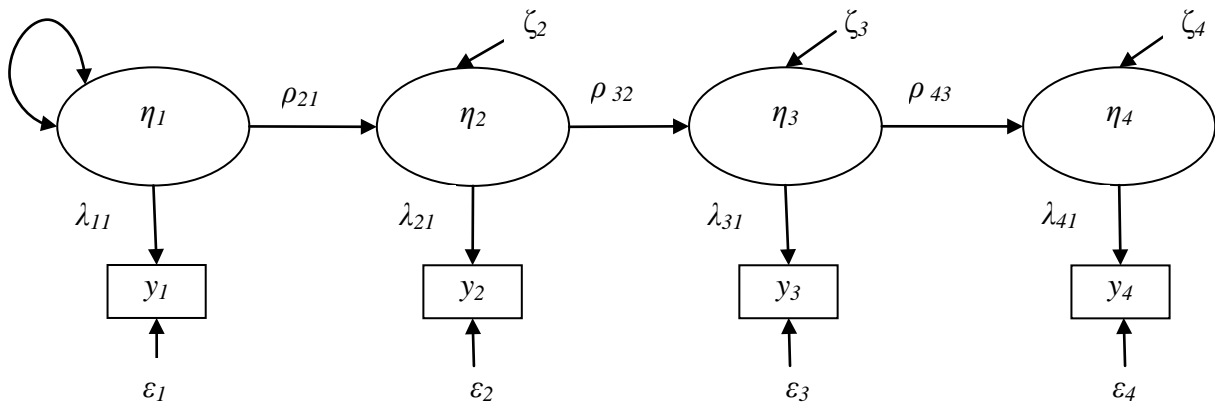
En autoregressiv modell er en longitudinell model der én variabel (for eksempel lesing) målt på et gitt tidspunkt blir kontrollert for den samme variabelen målt på et tidligere tidspunkt (ved bruk av regresjon). Autoregressorer benyttes for å unngå retningsproblematikk; avhengig variabel korrigeres for avhengig variabel på et tidligere tidspunkt. Autoregressive modeller analyserer endringer i rangering over tid, men sier ikke noe om hvordan det enkelte individs absolutte skårer endrer seg over tid. (Schlueter et al., 2007).

#### 3.3.1 Univariat latent autoregressiv modell

Den opprinnelige Markov Simplex modellen, som kun benyttet seg av observerte variabler er videreutviklet slik at man kan inkludere latente variabler i de autoregressive modellene.

”When measurement error of the observed variables is included in the model the simplex property shifts to the latent level. This model is known as the quasi simplex” (Bast & Reitsma, 1997). Slike autoregressive modeller baserer seg på at hvert latente begrep  $\eta_{it}$  målt

på tidspunkt  $t$  er en funksjon av verdien på tidspunkt  $t-1$  og tilfeldige feil. I tillegg trenger man en målemodell for å relatere de latente variablene til sine respektive indikatorer. Den autoregressive prosessen blir beskrevet av stabilitetskoeffisienter som reflekterer mengden av forandring i rekkefølgen mellom individer mellom to eller flere måletidspunkter (Schlueter et al., 2007).



**Figur 3** Univariat latent autoregressiv modell med målemodell

En univariat autoregressiv modell uttrykker de repeterte målingene av den latente variabelen  $\eta$  for individ  $i$  på tidspunkt  $t$  som en funksjon av et sett parametre:

$$\eta_{it} = \alpha_t + \rho_{t,t-1}\eta_{i,t-1} + \zeta_{it}$$

$\eta_{it}$  er verdien av den latente variabelen  $\eta$  for individ  $i$  på tidspunkt  $t$ ;  $\eta_{i,t-1}$  er verdien av den latente variabelen  $\eta$  for individ  $i$  på tidspunkt  $t-1$ ;  $\alpha_t$  er tidsspesifikt intercept;  $\rho_{t,t-1}$  er den autoregressive koeffisienten;  $\rho_{t,t-1}$  har samme verdi for alle casene i utvalget på et gitt tidspunkt og reflekterer styrken i sammenhengen mellom to tidspunkter;  $\zeta_{it}$  er residualen. Det forutsettes at  $\zeta_{it}$  har et gjennomsnitt på 0, ikke er autokorrelert og er ukorrelert med  $\eta_{i,t-1}$  (Schlueter et al., 2007; Bollen & Curran, 2006).



Målemodellen kan uttrykkes på denne måten:

$$y_{ikt} = \mu_{ikt} + \lambda_{kt}\eta_{ikt} + \varepsilon_{kit}$$

$y_{ikt}$  er den observerte verdien for en spesifikk ( $k = 1, 2, \dots, N$ ) indikator  $y$  for hvert individ på tidspunkt  $t$ ;  $\mu_{ikt}$  er intercept;  $\lambda_{kt}$  er faktorladningen som relaterer en spesifikk ( $k = 1, 2, \dots, N$ ) indikator til en latent faktor  $\eta_{ikt}$  og  $\varepsilon_{kit}$  er residualen (Schlueter et al., 2007).

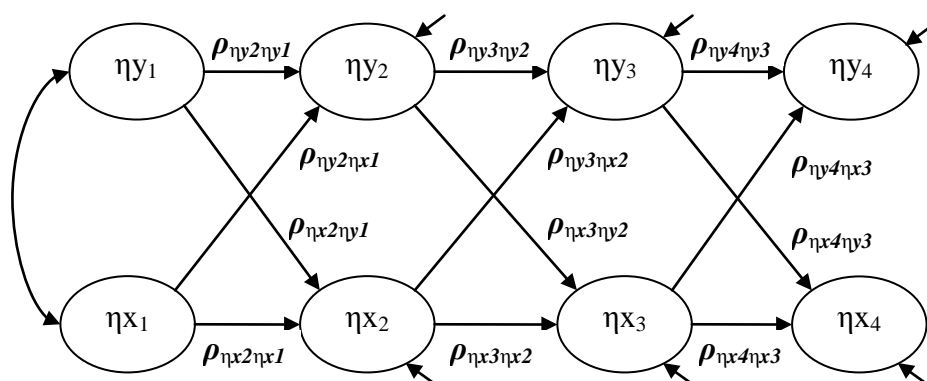
### 3.3.2 Bivariat latent autoregressiv modell

En bivariat autoregressiv modell med kryssladninger er en utvidelse av den univariate modellen. I en bivariat modell har man to sett av repeterte målinger over tid; for eksempel målinger av både ortografisk lesing og ordstaving på tidspunktene 2-5. En signifikant kryssladning mellom for eksempel ortografisk lesing på tidspunkt 3 og ordstaving på tidspunkt 2 tolkes da som at ordstaving på tidspunkt 2 predikerer ortografisk lesing på tidspunkt 3 utover det ortografisk lesing på tidspunkt 2 allerede bidrar med. "Due to the control of autoregression for each latent construct via the stability coefficients, the cross-lagged effects indicate the "pure" influence of each construct of interest" (Schlueter et al., 2007, s. 318-319). En illustrasjon av den bivariate modellen (med latente variabler) er gitt i figur 4. Målemodellen er ikke med i figur 4; den er illustrert i figur 3 (univariate autoregressiv modell).

$$\eta_{yit} = \alpha_{\eta y t} + \rho_{\eta y \eta y t-1} \eta_{yi,t-1} + \rho_{\eta y \eta x t-1} \eta_{xi,t-1} + \zeta_{\eta y it}$$

$$\eta_{xit} = \alpha_{\eta x t} + \rho_{\eta x \eta y t-1} \eta_{yi,t-1} + \rho_{\eta x \eta x t-1} \eta_{xi,t-1} + \zeta_{\eta x it}$$

$\alpha_{\eta y t}$  og  $\alpha_{\eta x t}$  representerer intercept til estimatene på hvert tidspunkt  $t$ .  $\rho_{\eta y \eta y t-1}$  og  $\rho_{\eta x \eta x t-1}$  er de autoregressive parametrene, og  $\rho_{\eta y \eta x t-1}$  og  $\rho_{\eta x \eta y t-1}$  er kryssladningene.  $\zeta_{\eta y it}$  og  $\zeta_{\eta x it}$  er residualene. Likningene er hentet fra Bollen & Curran (2004, s. 340), men er justert slik at de er med latente variabler og ikke observerte variabler som i den opprinnelige modellen ( $\eta y$  og  $\eta x$  istedenfor  $y$  og  $x$ ).  $\zeta$  benyttes som betegnelse på residualen istedenfor  $\varepsilon$  for å skille mellom residualen til de latente variablene i den autoregressive modellen og residualene i målemodellen.



Figuren er hentet fra Bollen & Curran, 2004, s. 341, men justert slik at den er med latente variabler og ikke observerte variabler som i den opprinnelige modellen.

**Figur 4** Bivariat latent autoregressiv modell med kryssladninger (uten målemodellen)

## 3.4 Data

Dataene jeg analyserer i oppgaven, er del av en større longitudinell studie om lese- og staveutvikling hos et representativt utvalg norske barn.

Redegjørelsen for dataene og variablene i undersøkelsen er hentet fra Lervåg, Bråten & Hulme (2009), Lervåg & Hulme (2009) og Lervåg & Hulme (2010).

### 3.4.1 Deltakere

233 norske førsteklasinger (123 jenter og 110 gutter) ble rekruttert til studien.

Gjennomsnittsalderen ved starten av studien var 6 år og 4 måneder (fra 5 år og 9 måneder til 6 år og 11 måneder). Alle hadde norsk som førstespråk, og ingen av barna hadde utviklingsvansker som for eksempel autisme eller alvorlige språkvansker eller sansetap som blind- eller døvhet som var oppdaget før skolestart. Barn fra alle skoler i Ski kommune ble invitert til å delta i studien. Av de foreldrene som sa ja til at barna kunne delta, ble det trukket et tilfeldig utvalg fra hver klasse. Fire barn ville ikke delta etter den første testingen, og to barn ble ekskludert på grunn av hørselsvansker. 228 barn deltok på alle testene ved første

testtidspunkt, og 192 barn deltok på alle testene ved alle testtidspunktene. Frafallet skyldes i hovedsak at barna flyttet ut fra kommunen.

### **3.4.2 Lese- og skriveopplæring**

I perioden 1997-2006, som dekker de årene denne studien pågikk, fulgte man i den norske skolen læreplanen L97. I disse årene startet man ikke med formell lese- og staveopplæring før i begynnelsen av andre klasse. I første klasse skulle man bli kjent med bokstaver og tall gjennom lek. Da den formelle opplæringen tok til i begynnelsen av andre klasse, ble barna hovedsaklig undervist ved hjelp av en fonetisk tilnærming.

### **3.4.3 Design og prosedyre**

Barna ble i ”moderundersøkelsen” testet på 5 ulike tidspunkter (T1-T5). I mine analyser benyttes bare deler av data fra denne undersøkelsen. Jeg benytter meg ikke av de kognitive variablene testet på T1. Mine data er kognitive variabler testet på T2 samt lese- og stavevariabler testet på T2-T5. Jeg har valgt å ikke gjøre om på betegnelsen på de ulike testtidspunktene, selv om jeg i mine analyser ikke benytter meg av variablene som ble målt på T1.

- T2 var i november og desember i andre klasse (rundt 3 måneder etter at de hadde startet med formell lese- og staveundervisning).
- T3 var i mai og juni (ved slutten av andre klasse).
- T4 var i november og desember i tredje klasse.
- T5 var i november og desember i fjerde klasse.

All testingen foregikk på skolen, og testene ble gitt i en bestemt rekkefølge.

### 3.4.4 Tester

#### **Fonologisk bevissthet målt ved T2:**

Fe\_2 er phoneme deletion. Denne testen hadde 24 items med økende vanskelighetsgrad. De åtte første itemene var multiple choice items med støtte av bilder. Her ble barnet spurt om hvilket av tre ord et ord ble endret til ved å fjerne et fonem fra enten begynnelsen eller slutten av ordet. I de neste åtte itemene ble barnet bedt om å lage et nytt ord ved å fjerne et fonem i begynnelsen eller slutten av et ord. Her var det ingen støtte av bilder. De siste åtte itemene bestod av 8 nonord som ble forandret til andre nonord. Maksimal skåre var 24.

Fs\_2 er phoneme segmentation. Denne testen hadde 24 items med økende vanskelighetsgrad. På de først 8 itemene ble barnet bedt om å telle antall fonemer i ordet ved hjelp av å telle på fingrene. Ordene bestod av to til fem fonemer. På de neste åtte itemene måtte barnet uttale alle fonemene i ord bestående av to til fem fonemer. De siste åtte itemene var lik som den forrige, men bestod av nonord. Maksimal skåre var 24.

Fi\_2 er phoneme isolation. Denne testen hadde 24 items med økende vanskelighetsgrad. De åtte første itemene var multiple choice items med støtte av bilder. Itemene ble presentert muntlig av testeren, og barnet ble bedt om å si hvilket av ordene som startet eller sluttet med et bestemt fonem. De neste åtte itemen var av typen “hva er den første lyden i ordet dag” og “hvilken lyd kommer etter /b/ i brann”? De åtte siste itemen var like som de sistnevnte, men bestod av nonord. Maksimal skåre var 24.

#### **Rapid Automatized Naming (RAN) målt ved T2:**

RAN ble målt ved bruk av to tester: ran\_n\_2 som er RAN numbers og ran\_l\_2 som er RAN letters. Her skulle barnet navngi 40 items (tall og bokstaver) arrangert i delvis tilfeldig rekkefølge på et ark: fem items gjentatt åtte ganger på fire linjer. Barnet navnga alle itemene så hurtig som mulig, og tiden det tok fra første til siste item ble målt. Før selve testen startet, forsikret man seg om at barnet kunne navnet på de itemene som var involvert i testene.

#### **Lesing målt ved T2-T5:**

Ort2\_2, ort2\_3, ort2\_4 og ort2\_5 er ortografisk lesing. Ortografisk lesing ble målt med en deltest fra STAS-batteriet som består av 77 rader med ord som høres like ut (som for eksempel kām-kom-komm). Bare ett av ordene er stavet riktig. Poengsummen ble regnet ut av

summen av alle markerte ord som var korrekte, minus antallet nonord som var galt markert, innenfor en tidsramme på 2 minutter.

NonRead2, NonRead3, NonRead4 og NonRead5 er nonordlesing. Nonordlesing ble målt med to nonordlesetester fra STAS-batteriet. Den ene bestod av 85 nonord med vanlige bokstavkonstellasjoner og den andre av 85 nonord med mindre vanlige bokstavkonstellasjoner. Barna måtte lese så mange nonord som mulig i løpet av 1 minutt. Variabelen er en samlev variabel av de to testene.

### **Staving målt ved T2-T5:**

Staving ble målt med to modifiserte tester fra STAS-batteriet: Staving av ord og staving av nonord. Ordene ble diktert av testadministratoren og barna måtte stave ordene på et papir. Barna fikk tid til å skrive ferdig et ord før det neste ble diktert. Det ble gitt 1 poeng for hvert ord eller nonord som ble stavet riktig.

Stav2, stav3, stav4 og stav5 er ordstaving. Denne ordstavetesten bestod av 47 ord med ulik vanskegrad med både høyfrekvente og lavfrekvente ord. Testen bestod av både konsistente ord (som ”diger”) og inkonsistente ord (som ”de”), samt ord som kan defineres av mer kompliserte fonem-grafemregler (som ”opp”) eller rot-morfemet (som ”også”).

Nonstav2, nonstav3, nonstav4 og nonstav5 er nonordstaving ved T2, T3, T4 og T5. Denne testen bestod av 25 nonord med varierende vanskegrad. 19 av ordene kunne staves ved å bruke ganske enkle fonem-grafemregler. Disse nonordene varierte i kompleksitet fra enkle CV nonord som ”ma” til mer komplekse CCVCV nonord som ”fnagy”. De siste 6 nonordene kunne staves korrekt bare dersom barna hadde kjennskap til mer komplekse fonem-grafemregler som inkluderte grafemer som for eksempel ”skj” eller ved å bruke regler om dobbelt konsonant etter et kort vokalfonem (som ”bemme”).

## 4 Resultater

Analysene er gjennomført ved hjelp av programmet Mplus, versjon 6.1 og SPSS, versjon 16.0 (den deskriptive statistikken). En inspeksjon av dataene viste at noen av variablene ikke var helt normalfordelte. Jeg har derfor benyttet meg av en estimeringsteknikk som er mer robust i forhold til non-normale data, nærmere bestemt Yuan-Bentler korreksjoner av ML-estimeringer med FIML-estimering av missing data. (MLR i Mplus, Muthen & Muthen, 1998-2010).

I alle tabeller og figurer betyr \*\*  $p < .01$  og \*  $p < .05$ .

### 4.1 Deskriptiv statistikk/korrelasjoner

Korrelasjonsmatrise for alle de observerte variablene er oppført i vedlegg 1.

Deskriptiv statistikk (gjennomsnitt, standardavvik, skjevhet og kurtosis) for de observerte kognitive variablene på tidspunkt 2 er oppført i tabell 1 sammen med korrelasjonene mellom dem. To av testene på fonologisk bevissthet viser moderate tegn til skjevhet og kurtosis. Fordelingene for phoneme segmentation og phoneme isolation er mer venstreskjeve (negativ skjevhet) og noe spissere enn normalfordelingen. De to RAN-testene avviker også moderat fra normalfordelingen ved å være spissere og høyreskjeve (positiv skjevhet). Alle variablene korrelerer signifikant med hverandre ( $p < .01$ ) Den sterkeste korrelasjonen er mellom de to RAN-testene (.72), mens de andre variablene korrelerer moderat til sterkt med hverandre (fra .34 til .56). RAN-verdiene er negative fordi denne testen gikk ut på at barna skulle navngi iteme så hurtig som mulig. Skalaen er altså motsatt på de to RAN-testene sammenlignet med resten av testene.

**Tabell 1** Deskriptiv statistikk og korrelasjoner for observerte kognitive variabler T2

	<b>M (SD)</b>	<b>Skjevhet</b>	<b>Kurtosis</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
<b>1.</b> Fe_2: Phoneme deletion	15.66 (5.85)	-.537	-.614	–	.46	.56	-.34	-.40
<b>2.</b> Fs_2: Phoneme segmentation	18.14 (6.06)	-1.473	1.105		–	.49	-.34	-.39
<b>3.</b> Fi_2: Phoneme isolation	20.24 (4.11)	-1.911	3.880			–	-.46	-.55
<b>4.</b> Ran_n_2: RAN numbers	38.51 (12.20)	1.507	2.894				–	.72
<b>5.</b> Ran_l_2: RAN letters	38.10 (14.81)	1.523	2.582					–

*Note.* Alle korrelasjonene er signifikante på 0.01-nivå.

Deskriptiv statistikk (gjennomsnitt, standardavvik, skjevhet og kurtosis) for de observerte lesevariablene (ortografisk lesing og nonordlesing) på tidspunktene 2-5 er oppført i tabell 2 sammen med korrelasjonene mellom dem. Alle variablene viser en adekvat fordeling. Det er økning i gjennomsnittsverdien for alle variablene fra T2 til T5. Korrelasjonen mellom alle variablene er moderate til høye på alle måletidspunkter. Korrelasjonen mellom ortografisk lesing og nonordlesing avtar fra T2 til T5. I de videre analysene er derfor ortografisk lesing og nonordlesing analysert hver for seg.

**Tabell 2** Deskriptiv statistikk og korrelasjoner for ortografisk lesing og nonordlesing T2-T5

	<b>M (SD)</b>	<b>Skjevhet</b>	<b>Kurtosis</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>
<b>1.</b> Ortografisk lesing T2	4.68 (4.14)	.960	.630	–	.57	.51	.43	.66	.53	.47	.42
<b>2.</b> Ortografisk lesing T3	9.53 (6.16)	.748	.630		–	.73	.60	.60	.65	.58	.54
<b>3.</b> Ortografisk lesing T4	14.09 (6.97)	.491	.454			–	.67	.51	.65	.58	.52
<b>4.</b> Ortografisk lesing T5	26.38 (9.53)	.308	-.004				–	.39	.47	.50	.57
<b>5.</b> Nonordlesing T2	25.83 (16.54)	.581	.246					–	.68	.64	.49
<b>6.</b> Nonordlesing T3	44.97 (18.81)	.795	1.071						–	.86	.76
<b>7.</b> Nonordlesing T4	53.61 (19.87)	.604	.698							–	.79
<b>8.</b> Nonordlesing T5	66.28 (23.94)	.795	.474								–

*Note.* Alle korrelasjonene er signifikante på 0.01-nivå.

Deskriptiv statistikk (gjennomsnitt, standardavvik, skjevhet og kurtosis) for de observerte stavevariablene (ordstaving og nonordstaving) på tidspunktene 2-5 er oppført i tabell 3 sammen med korrelasjonene mellom dem. Ordstaving viser en adekvat fordeling på alle måletidspunkter, mens nonordstaving viser moderate tegn på avvik fra normalfordelingen på tidspunkt 3 og 4. Det er økning i gjennomsnittsverdien for alle variablene fra T2 til T5. Korrelasjonen mellom alle variablene er moderate til høye på alle måletidspunkter. Korrelasjonen mellom ordstaving og nonordstaving avtar fra T2 til T5. I de videre analysene er derfor ordstaving og nonordstaving analysert hver for seg.

**Tabell 3** Deskriptiv statistikk og korrelasjoner for ordstaving og nonordstaving T2-T5

	M (SD)	Skjevhet	Kurtosis	1	2	3	4	5	6	7	8
1. Ordstaving T2	13.37 (7.09)	-.081	-.484	–	.72	.64	.64	.74	.56	.48	.47
2. Ordstaving T3	18.41 (6.66)	-.264	-.026		–	.75	.67	.66	.60	.48	.46
3. Ordstaving T4	22.38 (5.97)	-.259	.698			–	.71	.62	.54	.50	.44
4. Ordstaving T5	28.27 (6.16)	-.407	.035				–	.63	.62	.48	.54
5. Nonordstaving T2	11.37 (5.27)	-.619	-.545					–	.57	.43	.47
6. Nonordstaving T3	14.04 (3.76)	-1.165	1.378						–	.56	.49
7. Nonordstaving T4	15.30 (3.30)	-1.036	1.443							–	.49
8. Nonordstaving T5	16.93 (2.93)	-.588	.917								–

*Note.* Alle korrelasjonene er signifikante på 0.01-nivå.

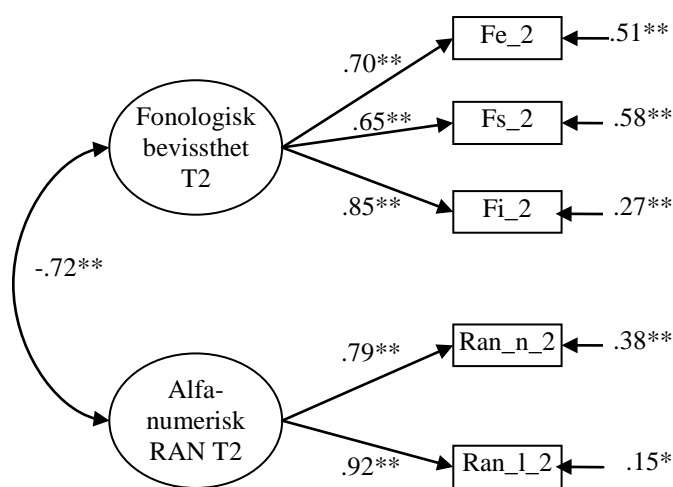
Korrelasjonen mellom nonordlesing og nonordstaving på de ulike tidspunktene er på henholdsvis .68, .43, .36 og .34. Det er altså sterk korrelasjon på T2, mens korrelasjonene på de tre andre tidspunktene er moderat.

Korrelasjonen mellom ortografisk lesing og ordstaving er på henholdsvis .60, .65, .64 og .60. Her er korrelasjonen høy og jevn hele veien.



## 4.2 Målemodeller

Jeg har estimert målemodellen for de latente kovariatene målt på T2. Resultatene er oppført i figur 5. Målemodellen har meget god modelltilpasning, og alle indikatorene har høye, standardiserte faktorladninger. De estimerte korrelasjonene mellom de latente kovariatene viser at det er høy korrelasjon mellom fonologisk bevissthet og alfanumerisk RAN målt på T2 (-.72). Videre i oppgaven er modeller der disse kovariatene er med fremstilt grafisk uten selve målemodellen. Dette for å gjøre fremstillingene enklere. Estimer for målemodellene for de fire univariate latente autoregressive modellene er oppført i tabell 4.



Modelltilpasning:  $YB\chi^2(4, N = 203) = 1.301, p < .861$ , RMSEA = .000 (90% CI = .000 - .056), CFI = 1.000, SRMR = .010

**Figur 5** Målemodell for latente kovariater T2

## 4.3 Latente vekstkurvemodeller

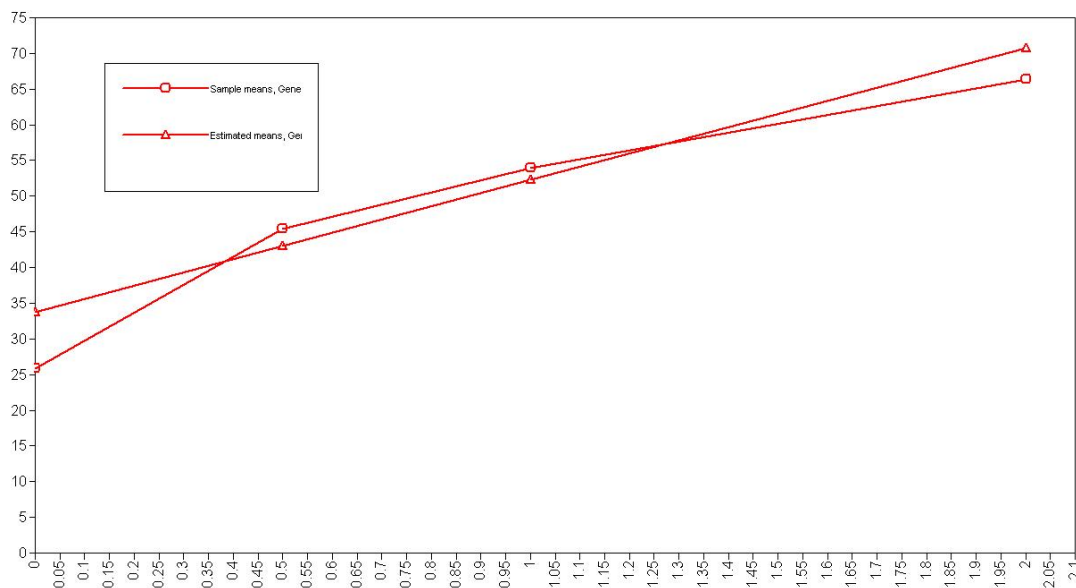
Det redegjøres først for estimeringen av ubetingede vekstkurvemodeller for de fire repeterte målingene (ortografisk lesing, ordstaving, nonordlesing og nonordstaving). Det ene målet med dette var å finne ut om veksten er lineær eller ikke, og hvordan en eventuelt ikke-lineær vekst best kunne modelleres. Det andre målet var å undersøke om det er signifikante variasjoner i de ulike vekstfaktorene. Jeg går så videre til estimering av bivariate latente

vekstkurvemodeller. Først analyseres nonordlesing og nonordstaving sammen, og deretter ortografisk lesing og ordstaving.

### 4.3.1 Ubetingede univariate latente vekstkurvemodeller

Det første jeg gjorde, var å undersøke om veksten til de ulike repeterte lese- og stavemålene var lineære eller ikke. Nedenfor følger en redegjørelse for hvordan jeg undersøkte dette for én av variablene, nærmere bestemt nonordlesing.

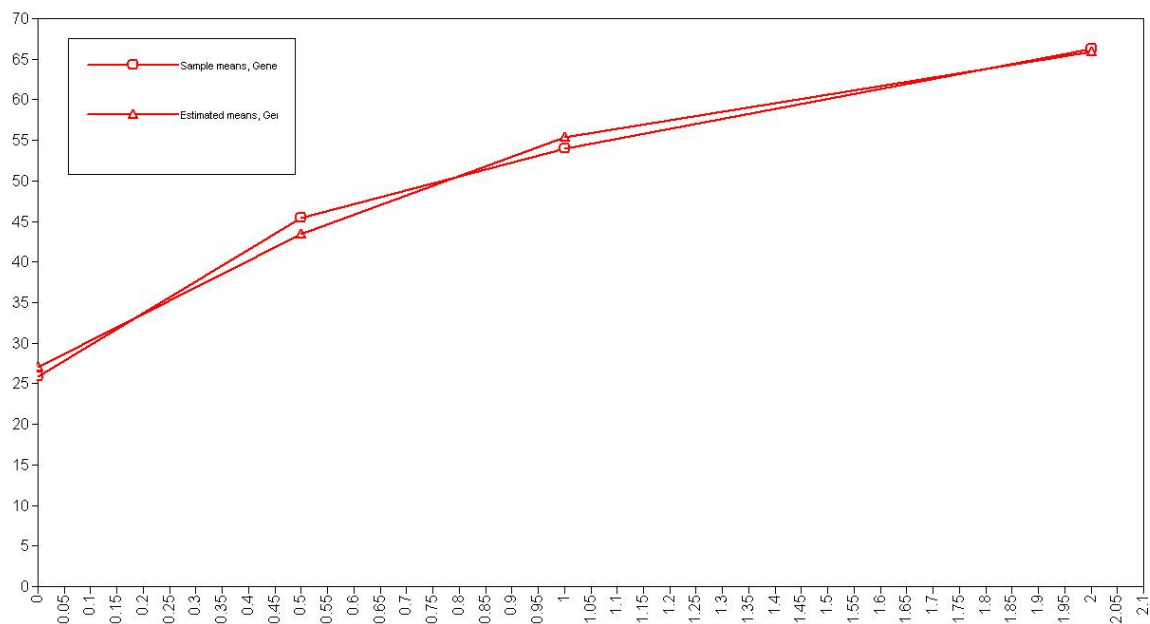
Jeg modellerte først den gjennomsnittlige veksten for nonordlesing estimert som lineær vekst. Modelltilpasningen var dårlig:  $YB\chi^2(5, N=208) = 161.714$ ,  $p = .000$ ,  $RMSEA = .388$  (90% CI = .338 - .441),  $CFI = .628$ ,  $SRMR = .162$ . Dette tyder på at veksten ikke er lineær. Den grafiske fremstillingen av det observerte gjennomsnittet i utvalget og det estimerte gjennomsnittet i modellen i figur 6 synliggjør at veksten ikke er lineær:



**Figur 6** Observert og estimert gjennomsnitt for nonordlesing, lineær modell

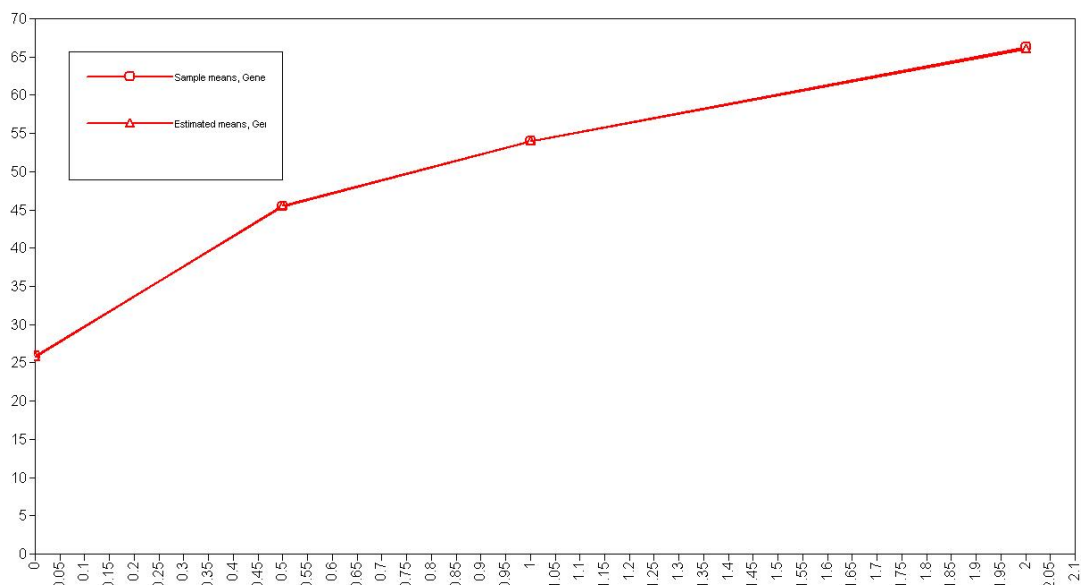
Deretter forsøkte jeg en modell med en kvadratisk faktor i tillegg til intercept- og slopefaktorene for å ta høyde for den ikke-lineære veksten.

Heller ikke denne modellen hadde så veldig god tilpasning, selv om den var bedre enn for den lineære modellen:  $YB\chi^2(4, N=208) = 45.957$ ,  $p = .000$ ,  $RMSEA = .225$  (90% CI = .169 - .285),  $CFI = .900$ ,  $SRMR = .094$ . Den grafiske fremstillingen i figur 7 synliggjør at heller ikke en slik modell samsvarer helt med dataene.



**Figur 7** Observert og estimert gjennomsnitt for nonordlesing, kvadratisk modell

Den siste modellen jeg prøvde ut, var en modell der de to midterste tidspunktene ble fritt estimert. Modelltilpasningen var god:  $YB\chi^2(3, N = 208) = 4.283$ ,  $p < .233$ ,  $RMSEA = .045$  (90% CI = .000 – .133),  $CFI = .997$ ,  $SRMR = .051$ . Den grafiske fremstillingen i figur 8 viser at denne modellen samsvarer godt med dataene. Verdien av Akaike Information Criteria (AIC) er 6411 for den kvadratiske modellen og 6371 for en ”completely latent” modell. For AIC-verdien gjelder det at ”minst er best”. Også dette viser at en ”completely latent” modell er den vekstmodellen som beskriver dataene best.



**Figur 8** Observert og estimert gjennomsnitt for nonordlesing, modell med to fritt estimerte slopefaktorer

For ortografisk lesing, ordstaving og nonordstaving ble de kvadratiske modellene vurdert til å ha god tilpasning, selv om verdien av RMSEA er noe høy for nonordstaving (.064). En completely latent modell ble derfor ikke testet ut for disse variablene.

Den valgte kvadratiske vekstmodellen for nonordstaving hadde en signifikant bedre tilpasning enn den lineære vekstmodellen:  $\Delta\chi^2(5) = 45.67$ ,  $p < .01$ . Den valgte modellen for nonordlesing (completely latent) hadde også en signifikant bedre tilpasning enn den lineære modellen:  $\Delta\chi^2(1) = 129.52$ ,  $p < .01$ . Ved bruk av MLR i estimeringen kan ikke  $\chi^2$ - verdien brukes direkte i differansetestingen mellom to modeller.  $\chi^2$ - differansetesten ble derfor gjennomført med korrigeringer.

For ortografisk lesing og ordstaving var verdien av  $MLR\chi^2$  mindre enn tilsvarende  $ML\chi^2$  for den lineære modellen. Jeg kunne dermed ikke benytte meg av den korrigerte differansetesten. Jeg har sett på AIC-verdien og ML-differansetesten isteden. For ortografisk lesing er AIC-verdien 4866 for den lineære modellen og 4856 for den kvadratiske modellen. En differansetest der  $ML\chi^2$  benyttes' viser at  $\Delta\chi^2(1) = 12.56$ ,  $p < .01$ . For ordstaving er AIC for henholdsvis den lineære og kvadratiske modellen 4798 og 4761. En differansetest der  $ML\chi^2$  benyttes' viser at  $\Delta\chi^2(3) = 43.48$ ,  $p < .01$ . Også for ortografisk lesing og ordstaving har altså en kvadratisk modell en signifikant bedre tilpasning enn en lineær modell.

En fullstendig oversikt over modelltilpasningen for de ulike univariate vekstmodellene er oppført i tabell 4.

**Tabell 4** Modelltilpasning for ubetingede latente vekstkurvmodeller

	N	Df	$YB\chi^2$	RMSEA med 90% CI	CFI	SRMR
<b>Ortografisk lesing lineær</b>	207	4	15.697 p = .004	.119 (.061-.183)	.951	.051
<b>Ortografisk lesing kvadratisk</b>	207	3	4.901 p = .179	.055 (.000-.140)	.992	.020
<b>Ordstaving Lineær</b>	204	5	42.769, p = 0.000	.192 (.142-.248)	.896	.092
<b>Ordstaving kvadratisk</b>	204	2	2.073 p = .355	.013 (.000-.140)	1.000	.025
<b>Nonordstaving lineær</b>	206	7	49.668 p = .000	.172 (.129-.219)	.763	.187
<b>Nonordstaving kvadratisk</b>	206	2	3.670, p = .160	.064 (.000-.166)	.991	.019
<b>Nonordlesing lineær</b>	208	5	161.714, p = .000	.388 (.338-.441)	.628	.162
<b>Nonordlesing kvadratisk</b>	208	4	45.957 p = .000	.225 (.169-.285)	.900	.094
<b>Nonordlesing completely latent</b>	208	3	4.283, p = .233	.045 (.000-.133)	.997	.051

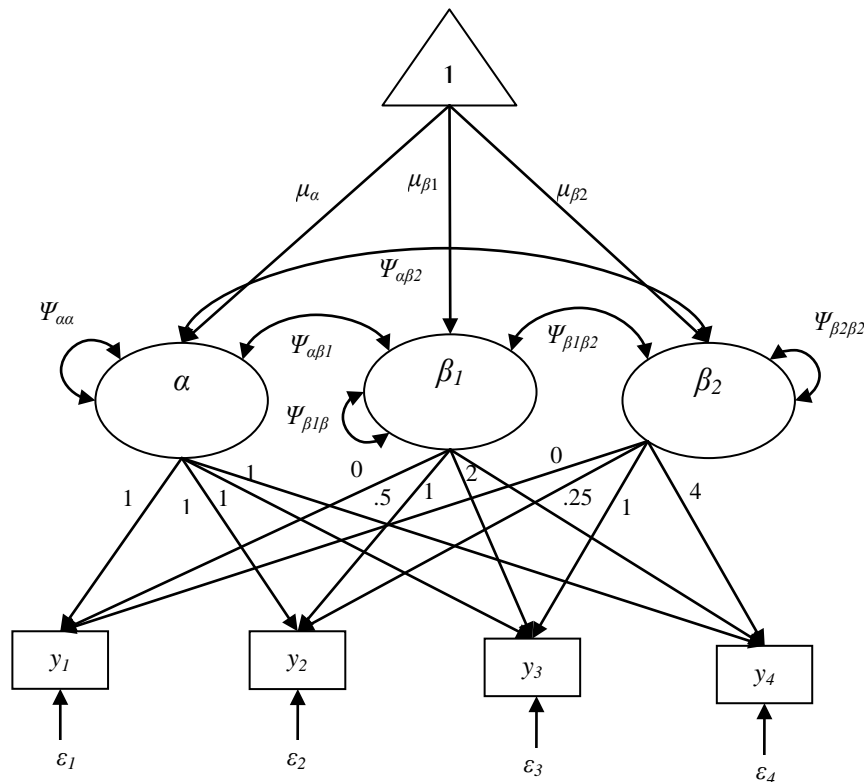
En kvadratisk modell ble altså vurdert til å beskrive dataene for ortografisk lesing, ordstaving og nonordstaving på en god måte. For nonordlesing ble en ”completely latent” modell vurdert til å beskrive dataene best. En videre analyse av de valgte univariate vekstmodellene, viser at det er signifikant varians i alle vekstfaktorene med unntak av slopefaktoren til nonordstaving (den er så vidt ikke signifikant,  $p = .061$ ) og den kvadratiske faktoren til ortografisk lesing. En

oppsummering av resultatene er oppført i tabell 5. De fire univariate vekstmodellene kommenteres hver for seg nedenfor. De ulike parametrene i en kvadratisk latent vekstkurvmodell illustreres i figur 9.

**Tabell 5** Ubetingede latente vekstkurvmodeller for ortografisk lesing, ordstaving, nonordstaving og nonordlesing

	Ortografisk lesing kvadratisk	Ordstaving kvadratisk	Nonordstaving kvadratisk	Nonordlesing Completely latent
<b>Varianser latente vekstfaktorer</b>				
$\psi_{\alpha\alpha}$	13.42**	42.42**	17.15**	215.77**
$\psi_{\beta1\beta1}$	12.57**	45.84**	19.98 <sup>NS</sup> (p=.061)	251.79**
$\psi_{\beta2\beta2}$	Satt til 0	9.32**	3.68*	–
<b>Kovarianser latente vekstfaktorer</b>				
$\psi_{\alpha\beta1}$	1.58 <sup>NS</sup> (.12 <sup>NS</sup> )	-22.16** (-.50**)	-13.56* (-.73**)	7.95 <sup>NS</sup> (.034 <sup>NS</sup> )
$\psi_{\alpha\beta2}$	–	7.64** (.11**)	4.40* (.55**)	–
$\psi_{\beta1\beta2}$	–	-19.31** (-.93**)	-8.01 <sup>NS</sup> (p=.057) (-.93**)	–
<b>Gjennomsnitt latente vekstfaktorer</b>				
$\mu_{\alpha}$	4.73**	13.39**	11.57**	25.76**
$\mu_{\beta1}$	8.33**	10.76**	5.20**	40.28**
$\mu_{\beta2}$	1.22**	-1.68**	-1.26**	–
<b>Residualvarians observerte variabler</b>				
var( $\varepsilon_1$ )	3.75* (.22*)	8.25* (.16**)	10.72** (.38**)	65.13** (.23**)
var( $\varepsilon_2$ )	17.19** (.49**)	10.97** (.26**)	4.97** (.36**)	55.54** (.16**)
var( $\varepsilon_3$ )	19.40** (.40**)	6.63** (.18**)	4.49** (.41**)	45.37** (.12**)
var( $\varepsilon_4$ )	19.81** (.22**)	Satt til 0	Satt til 0	142.65** (.23**)

*Note.* Ustandardiserte verdier med standardiserte verdier i parentes.



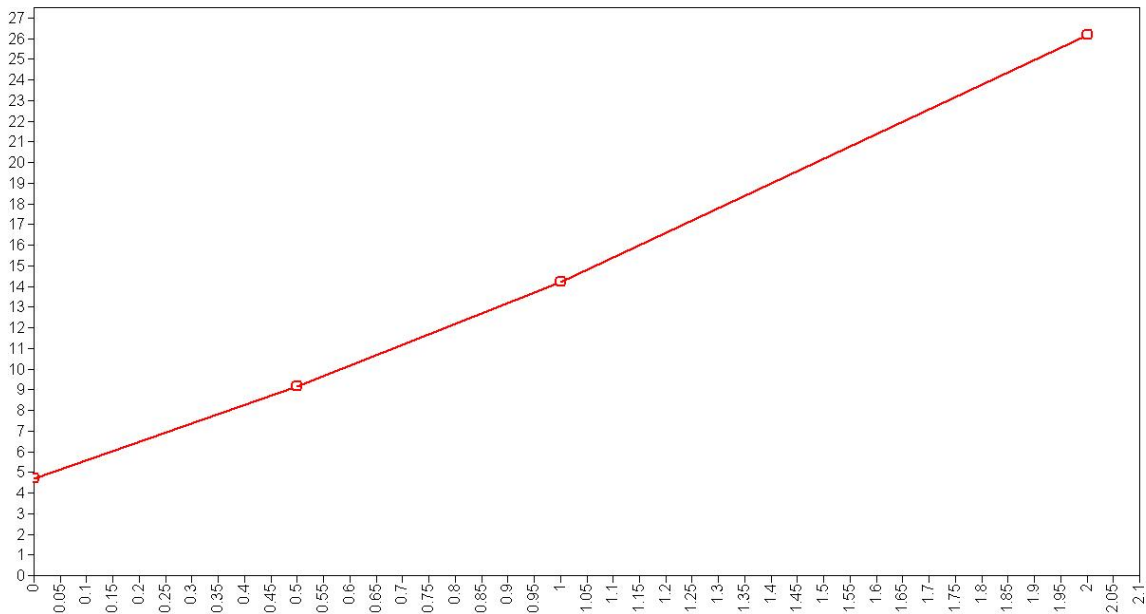
Figuren er delvis etter Bollen & Curran (2006) og delvis etter Preacher et al. (2008)

**Figur 9** Kvadratisk latent vekstkurvemodell

### Ortografisk lesing kvadratisk modell

Det var ikke signifikant varians i den kvadratiske faktoren. Denne ble derfor satt til 0. Modifikasjonsindeksen foreslo også å korrelere de observerte variablene på T3 og T4. Gjennomsnittet til interceptfaktoren ( $\mu_{\alpha}$ ) er 4.73. Dette betyr at barna i snitt skårer 4.73 poeng på det første måletidspunktet. Gjennomsnittet til den lineære vekstfaktoren (slopefaktoren  $\mu_{\beta_1}$ ) er 8.33. Dette betyr at barna i gjennomsnitt har en positiv lineær vekstkomponent i sine vekstkurver. Gjennomsnittet til den kvadratiske vekstfaktoren ( $\mu_{\beta_2}$ ) er 1.22. Dette betyr at kurven stiger brattere etter hvert som tiden går. I tillegg til disse signifikante "fixed effect components" er det også signifikante "random effect components" for to av de tre vekstkurvekomponentene. Det er individuelle variasjoner rundt gjennomsnittlig startverdi ( $\Psi_{\alpha\alpha} = 13.42$ ) og gjennomsnittlig lineær vekst ( $\Psi_{\beta_1\beta_1} = 12.57$ ), men ikke i den kvadratiske vekstkomponenten. Oppsummert viser resultatene fra den kvadratiske vekstkurvemodellen for ortografisk lesing at nivået av disse ferdighetene øker over tid, at veksten over tid blir

brattere og at det er individuelle forskjeller i startnivå og den lineære vekstkomponenten, men ikke i den kvadratiske komponenten. Det estimerte gjennomsnittet for vekstkurven er grafisk fremstilt i figur 10 og synliggjør at den gjennomsnittlige veksten blir brattere med tiden.



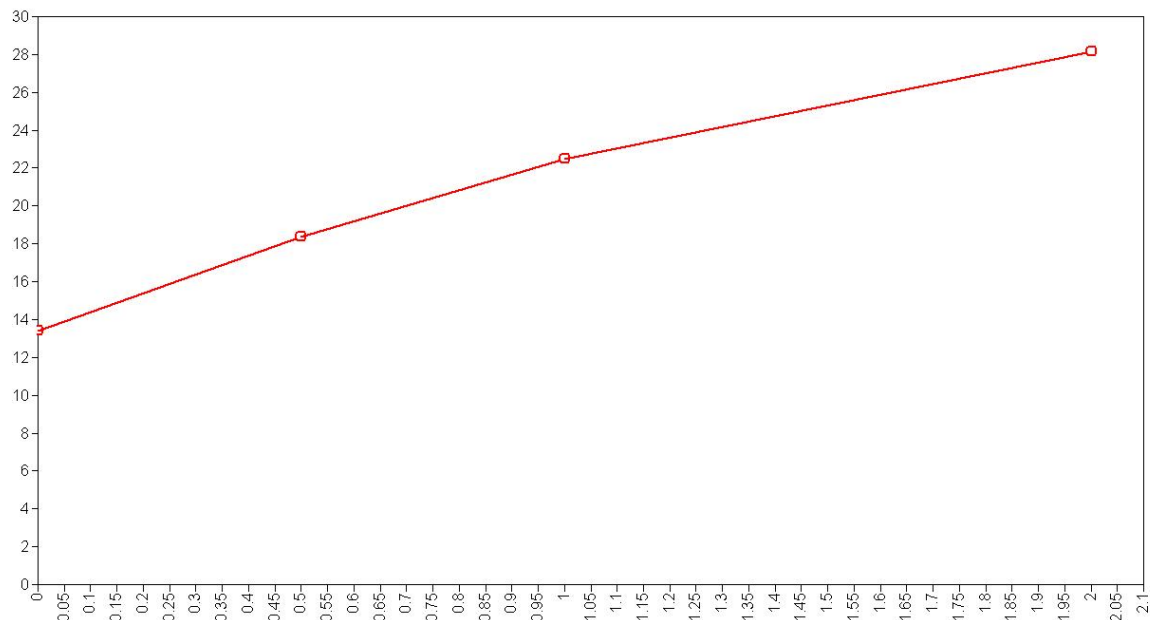
**Figur 10** Estimerte gjennomsnitt for ortografisk lesing, kvadratisk vekstkurvmodell

### Ordstaving (kvadratisk modell)

Korrelasjonen mellom slopefaktoren og den kvadratiske faktoren var større enn 1. Det viste seg at residualvariansen til den observerte variabelen på T5 ikke var signifikant. Denne ble derfor satt til 0. Gjennomsnittet til interceptfaktoren ( $\mu_{\alpha}$ ) er 13.39. Dette betyr at barna i snitt skårer 13.39 poeng på det første måletidspunktet. Gjennomsnittet til den lineære vekstfaktoren (slopefaktoren  $\mu_{\beta1}$ ) er 10.76. Dette betyr at barna i gjennomsnitt har en positiv lineær vekstkomponent i sine vekstkurver. Gjennomsnittet til den kvadratiske vekstfaktoren ( $\mu_{\beta2}$ ) er -1.68. Dette betyr at kurven stiger mindre bratt etter hvert som tiden går. I tillegg til disse signifikante "fixed effect components" er det også signifikante "random effect components" for alle de tre vekstkurvekomponentene. Det er individuelle variasjoner rundt gjennomsnittlig startverdi ( $\Psi_{\alpha\alpha} = 42.42$ ), i den gjennomsnittlig lineær veksten ( $\Psi_{\beta1\beta1} = 45.84$ ), og i den kvadratiske vekstkomponenten ( $\Psi_{\beta2\beta2} = 9.32$ ). Oppsummert viser resultatene fra den kvadratiske vekstkurvmodellen for ordstaving at nivået av disse ferdighetene øker over tid, men at veksten etter hvert blir mindre bratt. Det er individuelle forskjeller i både startnivå,



den lineære vekstkomponenten og den kvadratiske vekstkomponenten. Det estimerte gjennomsnittet for vekstkurven er grafisk fremstilt i figur 11 og synliggjør at den gjennomsnittlige veksten blir mindre bratt med tiden.



**Figur 11** Estimerte gjennomsnitt for ordstaving, kvadratisk vekstkurvmodell

### Nonordstaving (kvadratisk modell)

Residualvariansen til den observerte variabelen på T5 var ikke signifikant, og ble derfor satt til 0. Gjennomsnittet til interceptfaktoren ( $\mu_{\alpha}$ ) er 11.57. Dette betyr at barna i snitt skårer 11.57 poeng på det første måletidspunktet. Gjennomsnittet til den lineære vekstfaktoren (slopefaktoren  $\mu_{\beta 1}$ ) er 5.20. Dette betyr at barna i gjennomsnitt har en positiv lineær vekstkomponent i sine vekstkurver. Gjennomsnittet til den kvadratiske vekstfaktoren ( $\mu_{\beta 2}$ ) er -1.26. Dette betyr at kurven stiger mindre bratt etter hvert som tiden går. I tillegg til disse signifikante "fixed effect components" er det også signifikante "random effect components" for alle de tre vekstkurvekomponentene. Det er individuelle variasjoner rundt gjennomsnittlig startverdi ( $\Psi_{\alpha\alpha} = 17.15$ ) og i den kvadratiske vekstkomponenten ( $\Psi_{\beta 2\beta 2} = 3.68$ ). Variansen i den lineære vekstkomponenten er akkurat ikke signifikant ( $\Psi_{\beta 1\beta 1} = 19.98$ ,  $p = .061$ ). Oppsummert viser resultatene fra den kvadratiske vekstkurvmodellen for nonordstaving at nivået av disse ferdighetene øker over tid, men at veksten etter hvert blir mindre bratt. Det er individuelle forskjeller i startnivå og i den kvadratiske vekstkomponenten.

### Nonordlesing (completely latent modell)

Gjennomsnittet til interceptfaktoren ( $\mu_{\alpha}$ ) er 25.76. Dette betyr at barna i snitt skårer 25.76 poeng på det første måletidspunktet. Det er signifikant gjennomsnittlig vekst i nonordlesing mellom T2 og T5 (40.28). Dette betyr at barna i gjennomsnitt skårer 40.28 poeng mer på T5 enn på T2. Jeg har modellert faktorladningene fra slope til nonread2 lik 0 og nonread5 lik 1. Nonread3 og nonread4 er fritt estimert. Ved å kode tid på denne måten, blir de estimerte standardiserte faktorladningene for slope på henholdsvis .49 og .70. Disse tolkes som at 49% av den gjennomsnittlige veksten skjedde mellom T2 og T3, og 70% av den gjennomsnittlige veksten skjedde mellom T2 og T4. Den gjennomsnittlige veksten er altså størst i starten og flater ut etter hvert. Det er signifikant variasjon rundt gjennomsnittlig startverdi ( $\Psi_{\alpha\alpha} = 215.77$ ) og gjennomsnittlig vekst ( $\Psi_{\beta_1\beta_1} = 251.79$ ).

Estimatet av residualvariansen til de observerte variablene (altså de gjentatte målingene) sier noe om variansen i de gjentatte målingene som ikke forklares av den underliggende vekstprosessen (Bollen & Curran, 2006).

Jeg finner at det er signifikant varians i alle faktorene med unntak av slopefaktoren for nonordstaving og den kvadratiske faktoren for ortografisk lesing. Dette betyr at det er interessant å undersøke mulige forklaringer av disse variasjonene. Dette er tema i neste avsnitt der jeg tar for meg bivariate latente vekstkurvmodeller og også trekker inn fonologisk bevissthet og alfanumerisk RAN som kovariater.

#### 4.3.2 Bivariate latente vekstkurvmodeller

For å teste ut forholdet mellom utviklingen til de ulike lese- og stavevariablene i undersøkelsen, har jeg estimert bivariate latente vekstkurvmodeller. Jeg starter med bivariate modeller for nonordlesing og nonordstaving og fortsetter med bivariate modeller for ortografisk lesing og ordstaving.

Modellene mine er estimert med regresjoner mellom interceptfaktoren og slopefaktoren og mellom slopefaktoren og den kvadratiske faktoren. Dette gjør det mulig å vurdere om man kan si noe om variasjoner i veksten (slope) etter å ha kontrollert for ferdigheter ved første måletidspunkt (intercept), og variasjoner i den ikke-lineære veksten (kvadratisk) etter å ha

kontrollert for den lineære veksten (slope). Å ikke ta hensyn til forholdet mellom de ulike vekstfaktorene kan føre til mistolkninger. Dersom interceptfaktoren er sterkt relatert til både senere vekst og et kovariat, kan det være at kovariatet predikerer vekst. Ikke nødvendigvis fordi det er direkte relatert til vekst, men fordi kovariatet er relatert til intercept og dermed indirekte relatert til veksten. I en slik situasjon er kovariatets evne til å predikere vekst konfundert med kovariatets korrelasjon med intercept. (Lervåg & Hulme, 2010). Spørsmålet er altså om kovariatet predikerer videre vekst når det kontrolleres for ferdigheter på starttidspunktet.

Jeg har i figurene 12a og 13a tatt med hele vekstkurvemodellen for å vise hvordan jeg har kodet tiden på slopefaktoren og den kvadratiske faktoren. Det er et halvt år mellom T2-T3 og T3-T4 og ett år mellom T4 og T5. T2 er kodet som 0, og er dermed intercept i modellen. Intercept reflekterer det gjennomsnittlige nivået av henholdsvis ortografisk lesing og ordstaving på T2 (det første måletidspunktet). Siden det er et halvt år mellom T2 og T3 og mellom T3 og T4, er faktorladningene fra slope til de observerte variablene på T3 og T4 kodet som henholdsvis .5 og 1. Det er ett år mellom T4 og T5, og faktorladningen fra slope til den observerte variabelen på T5 er derfor kodet som 2. Faktorladningene for den kvadratiske faktoren er kvadrerte verdier av faktorladningene fra slopefaktoren.

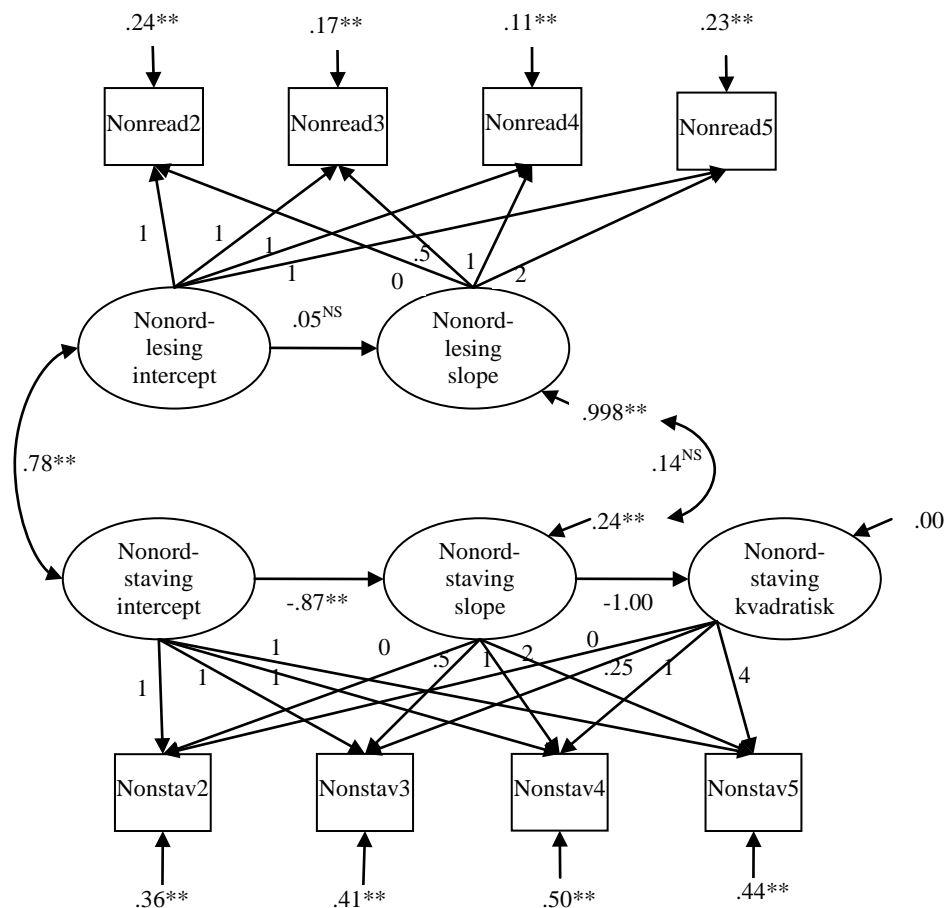
Modellene 12b, 12c og 13b har en mer forenklet grafisk fremstilling, der kodingen av tid og residualvariansen til de observerte variablene ikke er tatt med i den grafiske fremstillingen av modellene.

### **Bivariat latent vekstkurvemodell nonordlesing og nonordstaving uten kryssregresjoner**

Residualvariansen til den kvadratiske faktoren for nonordstaving var negativ, men ikke signifikant, og ble derfor satt til 0. Det er tatt høyde for kovarians mellom de observerte lese- og stavevariablene målt på samme tidspunkt. Modelltilpasningen er god.

Intercept til nonordstaving forklarer 76% av variansen i slope, og slope forklarer fullt ut hvordan veksten etter hvert avtar. Negativt fortegn betyr at det er de som starter lavest som vokser mest i starten, og at det er disse som etter hvert også har den sterkeste avtagende veksten. For nonordlesing er det ingen signifikant sammenheng mellom intercept og slope.

Det er altså ingen sammenheng mellom hvilke ferdigheter man har på T2 (intercept) og den videre veksten nonordlesing. Det er sterk korrelasjon mellom interceptfaktorene til nonordstaving og nonordlesing (.78). Det er ingen signifikant sammenheng mellom residualene til slopefaktorene, og dette kan tolkes som at den lineære veksten til nonordstaving og veksten til nonordlesing ikke har noe til felles når det kontrolleres for nivået av de respektive ferdighetene på T2.



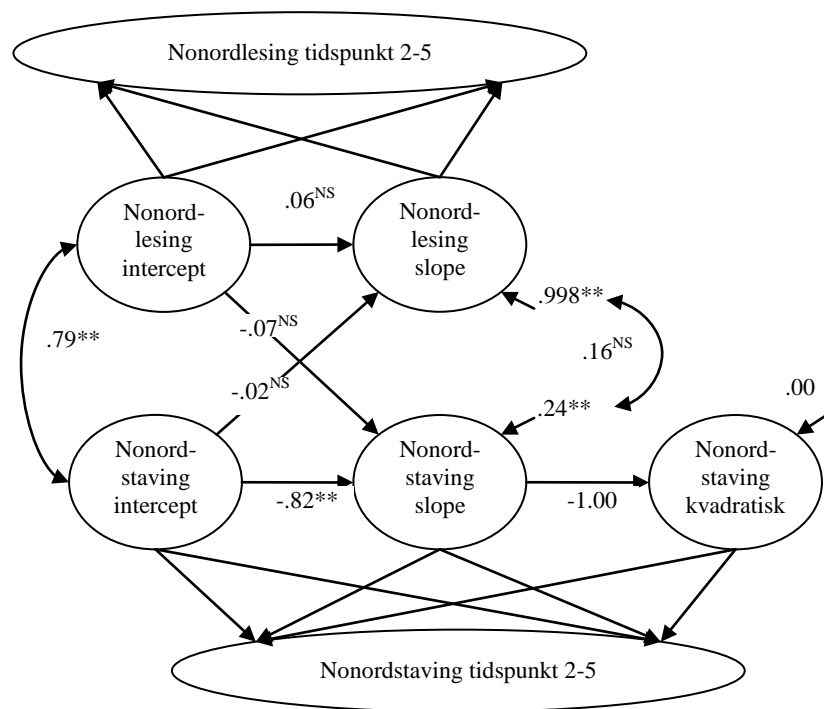
Modelltilpasning:  $YB\chi^2(19, N = 209) = 17.930, p < .527, RMSEA = .000$  (90% CI = .000 - .057), CFI = 1.000, SRMR = .045

**Figur 12a** Bivariat latent vekstkurvemodell nonordlesing og nonordstaving uten kryssregresjoner

## Bivariat latent vekstkurvemodell nonordlesing og nonordstaving med kryssregresjoner

For å forenkle den grafiske fremstillingen er figurene 12b og 12c uten de repeterte observerte variablene i figur 12a. Ingen av kryssregresjonene blir signifikante. Jeg finner altså ingen sammenheng mellom nonordstaving på starttidspunktet og veksten i nonordlesing, eller mellom nonordlesing på starttidspunktet og den lineære veksten i nonordstaving.

De øvrige resultatene i denne modellen sammenfaller i stor grad med resultatene i modellen uten kryssregresjoner (figur 12a). I neste modell (figur 12c) trekker jeg inn fonologisk bevissthet og alfanumerisk RAN som kovariater. Fordi ingen av kryssregresjonene var signifikante, er disse ikke med i den neste modellen med kovariater.



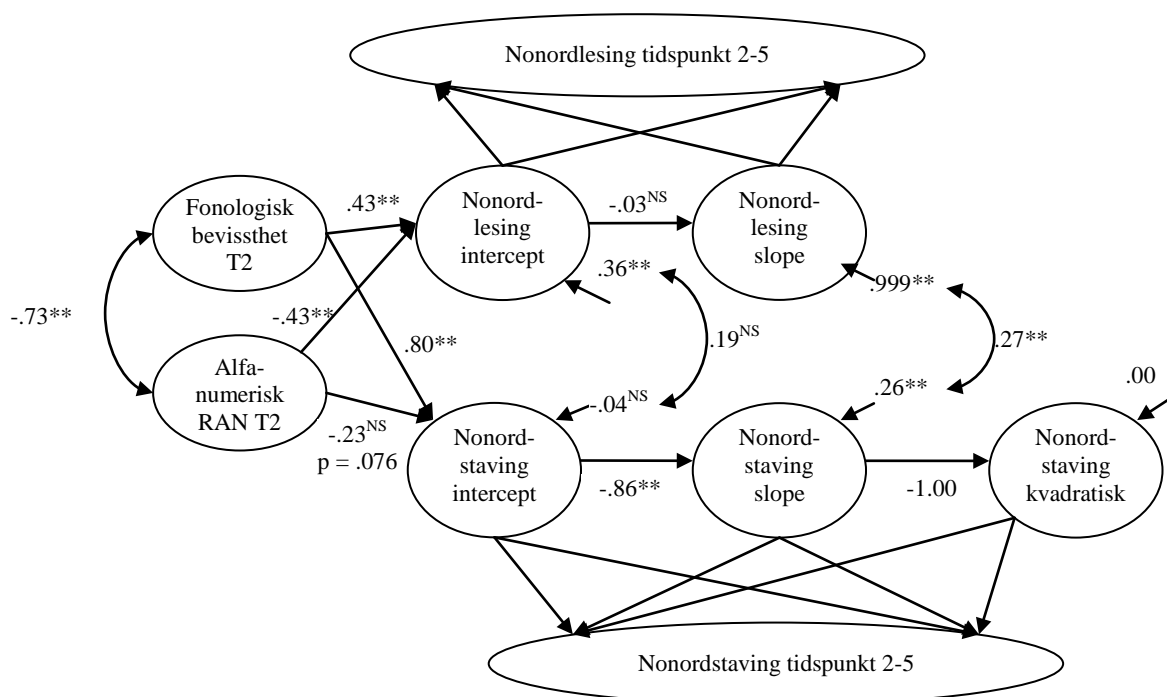
Modelltilpasning:  $YB\chi^2(17, N = 209) = 17.746, p < .405$ , RMSEA = .014 (90% CI = .000 - .065), CFI = .999, SRMR = .045

**Figur 12b** Bivariat latent vekstkurvemodell nonordlesing og nonordstaving med kryssregresjoner

## Bivariat latent vekstkurvemodell nonordlesing og nonordstaving med kovariater

Resultatene viser at fonologisk bevissthet predikerer variasjon i nonordstaving (.80), og variasjon i lesing av nonord (.43) på T2. Alfa-numerisk RAN predikerer variasjon i lesing av nonord (-.43) men ikke variasjon i nonordstaving når det kontrolleres for effekten av fonologisk bevissthet. I en modell der alfa-numerisk RAN er inne som eneste kovariat, predikerer dette kovariatet variasjoner i nonordstaving meget sterkt (-.82). Resultatene viser videre at fonologisk bevissthet og alfa-numerisk RAN til sammen forklarer henholdsvis all variasjonen i nonordstaving og 64% av variasjonen i nonordlesing på T2 (intercept).

I modellen der både alfa-numerisk RAN og fonologisk bevissthet er inne som kovariater, er det ikke lenger en signifikant sammenheng mellom de to interceptfaktorene. Den sterke sammenhengen jeg fant mellom de to interceptfaktorene i modell 12a og 12b forklares altså fullt ut av de to kovariatene. Modelltilpasningen er også her god.



Modelltilpasning:  $YB\chi^2(59, N = 209) = 58.853, p < .481, RMSEA = .000$  (90% CI = .000 - .042),  
CFI = 1.000, SRMR = .042

**Figur 12c** Bivariat latent vekstkurvemodell nonordlesing og nonordstaving med kovariater

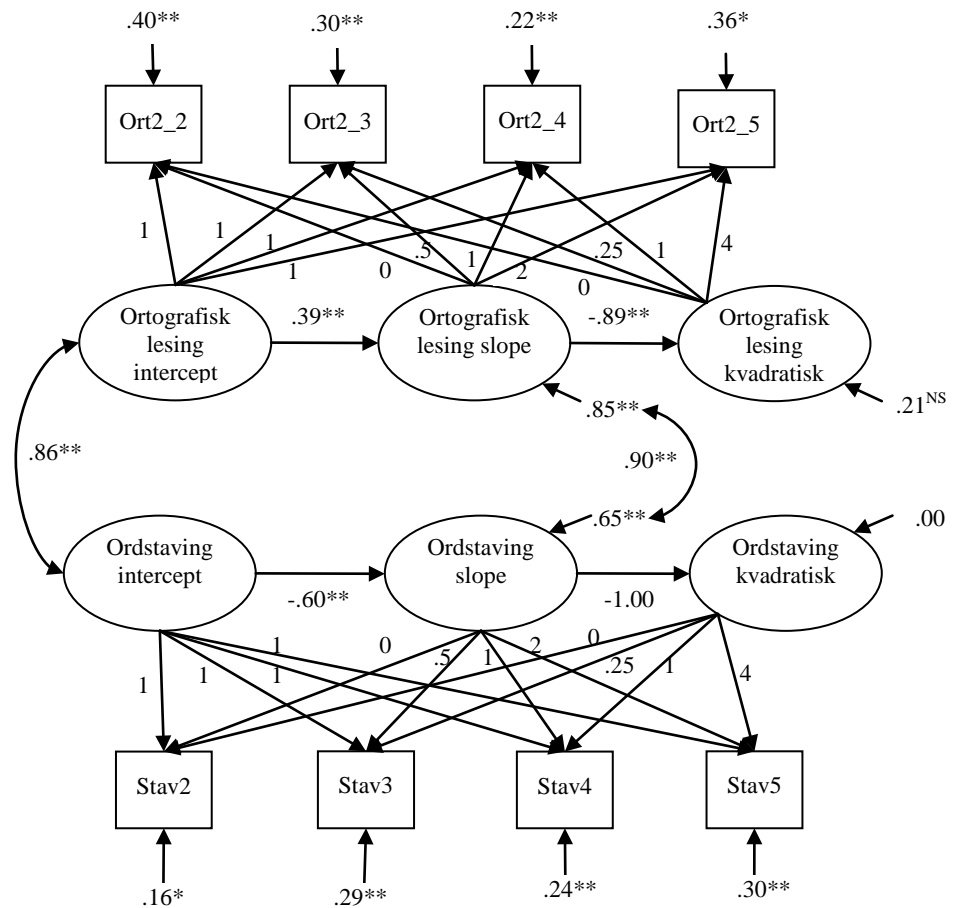
## **Bivariat latent vekstkurvmodell ortografisk lesing og ordstaving uten kryssregresjoner**

Residualvariansen til den kvadratiske faktoren for ordstaving var negativ, men ikke signifikant, og ble derfor satt til 0. Det er tatt høyde for kovarians mellom de observerte lese- og stovevariablene målt på samme tidspunkt. Modelltilpasningen er god.

Intercept til ordstaving forklarer 35% av variansen i slope, og slope forklarer fullt ut hvordan veksten etter hvert avtar. Negativt fortegn betyr at det er de som starter lavest som vokser mest i starten, og at det er disse som etter hvert også har den sterkeste avtagende veksten. Intercept til ortografisk lesing forklarer 15% av variansen i slope. Her er regresjonen positiv, og tolkes som at de som starter høyest vokser mest i starten. Men denne sammenhengen er ikke veldig sterk. Slope forklarer fullt ut hvordan de barna som vokser mest i starten etter hvert vokser mindre bratt.

Den høye positive korrelasjonen mellom interceptfaktoren til ortografisk lesing og interceptfaktoren til ordstaving (.86) reflekterer at i gjennomsnitt tenderer de individuelle verdiene som er over gjennomsnittet på ortografisk lesing på første måletidspunkt til å være over gjennomsnittet også på ordstaving på det første måletidspunktet. Det er altså en sterk sammenheng mellom ferdigheter i ortografisk lesing og ordstaving på T2. Det er også en sterk sammenheng mellom residualene til slopefaktorene (.90), og disse residualene er ganske store. Dette betyr at den lineære veksten til ortografisk lesing og ordstaving har noe til felles som ikke forklares av de respektive interceptfaktorene.

Som for nonordlesing og nonordstaving, testet jeg så ut en modell med kryssregresjoner. Ingen av disse var signifikante. I den neste modellen (med kovariater) er disse kryssregresjonene derfor ikke med.



Modelltilpasning:  $YB\chi^2(18, N = 207) = 30.575, p < .032, RMSEA = .058$  (90% CI = .017 - .093), CFI = .986, SRMR = .033

**Figur 13a** Bivariat latent vekstkurvemodell ortografisk lesing og ordstaving uten kryssregresjoner

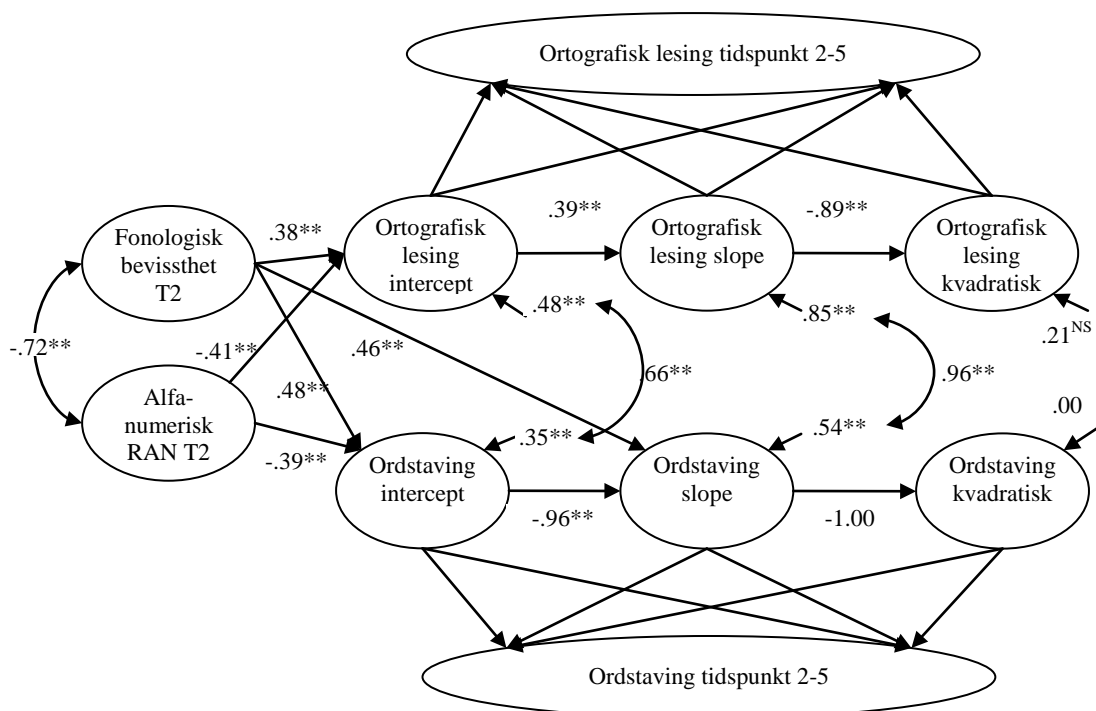
### Bivariat latent vekstkurvemodell ortografisk lesing og ordstaving med kovariater

Resultatene viser at fonologisk bevissthet predikerer variasjon i ordstaving (.48) og i ortografisk lesing (.38) på T2. Også alfanumerisk RAN predikerer variasjon i både ortografisk lesing (-.41) og ordstaving (-.39). Fonologisk bevissthet bidrar dessuten til å predikere videre lineær vekst (slope) i ordstaving (.45).

Resultatene viser videre at fonologisk bevissthet og alfanumerisk RAN til sammen forklarer henholdsvis 52% av variasjonen i ortografisk lesing på T2 (intercept) og 65% av variasjonen i



Det er høy positiv korrelasjon mellom residualene til interceptfaktoren for ortografisk lesing og interceptfaktoren for ordstaving (.66). Dette reflekterer at ortografisk lesing og ordstaving på starttidspunktet har noe til felles utover det som predikeres av fonologisk bevissthet og alfanumerisk RAN. Det er også en sterk sammenheng mellom residualene til slopefaktorene (.96), og disse residualene (altså uforklart varians) er ganske store. Dette betyr at også den lineære veksten til ortografisk lesing og ordstaving har noe til felles som ikke forklares av de respektive interceptfaktorene eller kovariatene.



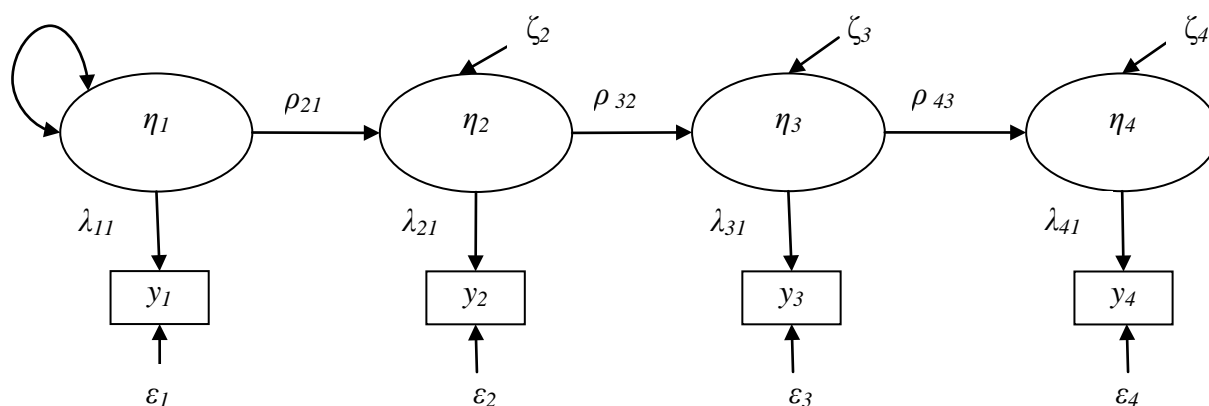
**Figur 13b** Bivariat latent vekstkurvemodell ortografisk lesing og ordstaving med kovariater

## 4.4 Autoregressive modeller

Jeg estimerte først univariate modeller for de nonordlesing, nonordstaving, ortografisk lesing og ordstaving. Resultatene er oppført i tabell 6. Så estimerte jeg bivariate modeller med kryssladninger, først uten og så med fonologisk bevissthet og alfanumerisk RAN som kovariater. Kombinasjonen av lese- og stavevariabler i de bivariate autoregressive modellene er de samme som for de latente vekstkurvemodellene; først nonordlesing og nonordstaving og så ortografisk lesing og ordstaving.

### 4.4.1 Univariate latente autoregressive modeller

De latente begrepene for henholdsvis ortografisk lesing, ordstaving, nonordlesing og nonordstaving på hvert måletidspunkt består av kun én observert variabel. For at modellen skulle identifiseres, ble residualvariansen til de observerte variablene satt til å være lik for alle måletidspunkter, med unntak av modellen for ortografisk lesing. Her ble residualvariansen på henholdsvis måletidspunktene 1 & 2 og 3 & 4 satt til å være like for at modellen skulle konvergere. Resultatene viser at det er høy stabilitet for alle lese- og stavevariablene gjennom hele perioden (høy stabilitetskoeffisient): De standardiserte estimatene er på henholdsvis .82, .88 og .75 for ortografisk lesing, .87, .94 og .92 for ordstaving, .76, .93 og .85 for nonordlesing og .77, .92 og 1.00 for nonordstaving. En så høy stabilitet mellom måletidspunktene betyr at de autoregressive effektene er sterke. Estimer for de fire univariate modellene er oppført i tabell 6. Modelltilpasningen synliggjøres i tabell 7. Modelltilpasningen er god for alle de fire univariate modellene, med unntak av RMSEA-verdien for nonordlesing som er noe høy (.071).



**Figur 14** Univariat latent autoregressiv modell med målemodell

**Tabell 6** Estimerte univariate latente autoregressive modeller

Parametre	Ortografisk lesing	Ordstaving	Nonordlesing	Nonordstaving
$\rho_{21}$	1.40** (.82**)	.81** (.87**)	.89** (.76**)	.50** (.77**)
$\rho_{32}$	1.00** (.88**)	.82** (.94**)	.99** (.93**)	.74** (.92**)
$\rho_{43}$	1.08** (.75**)	.96** (.92**)	1.03** (.85**)	.83** (1.00**)
var ( $\eta_1$ )	10.51** (1.00)	42.13** (1.00)	244.65** (1.00)	23.10** (1.00)
var ( $\zeta_2$ )	10.34 <sup>NS</sup> (.34 <sup>NS</sup> )	8.79** (.24**)	139.00** (.42**)	3.88** (.41**)
var ( $\zeta_3$ )	8.77* (.22*)	3.34* (.12*)	51.80** (.14**)	.93 <sup>NS</sup> (.15 <sup>NS</sup> )
Var ( $\zeta_4$ )	35.83** (.44**)	4.96 <sup>NS</sup> (.16 <sup>NS</sup> )	153.26** (.28**)	Satt til 0 pga negativ, men ikke signifikant residualvarians
$\lambda_{11}$	1.00 (.79**)	1.00 (.92**)	1.00 (.95**)	1.00 (.91**)
$\lambda_{21}$	1.00 (.91**)	1.00 (.91**)	1.00 (.96**)	1.00 (.82**)
$\lambda_{31}$	1.00 (.91**)	1.00 (.88**)	1.00 (.97**)	1.00 (.76**)
$\lambda_{41}$	1.00 (.95**)	1.00 (.89**)	1.00 (.98**)	1.00 (.69**)
var ( $\varepsilon_1$ )	6.50** (.38**)	7.91** (.16**)	26.33** (.10**)	4.59** (.17**)
var ( $\varepsilon_2$ )	6.50** (.17**)	7.91** (.18**)	26.33** (.07**)	4.59** (.32**)
var ( $\varepsilon_3$ )	8.18** (.17)	7.91** (.22**)	26.33** (.07)	4.59** (.43**)
var ( $\varepsilon_4$ )	8.18** (.09)	7.91** (.21**)	26.33** (.05)	4.59** (.52**)

*Note.* Ustandardiserte verdier med standardiserte verdier i parentes.  $\rho$  er stabilitetskoeffisienten mellom to måletidspunkter (de standardiserte verdiene); var( $\zeta_1$ ) er variansen til den latente variabelen på første måletidspunkt; var( $\zeta_2$ ) til var( $\zeta_4$ ) er residualvariansen til de latente variablene fra tidspunkt 2 til tidspunkt 4;  $\lambda_{11}$  til  $\lambda_{41}$  er faktorladninger; var( $\varepsilon_1$ ) til var( $\varepsilon_4$ ) er residualvariansen til de observerte variablene.

**Tabell 7** Modelltilpasning univariate latente autoregressive modeller

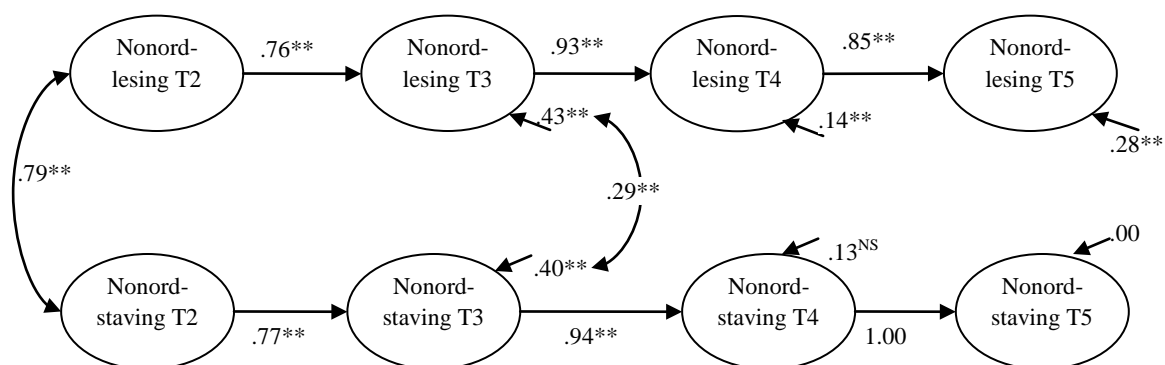
	<b>N</b>	<b>Df</b>	<b>YB<math>\chi^2</math></b>	<b>RMSEA med 90% CI</b>	<b>CFI</b>	<b>SRMR</b>
Ortografisk lesing	207	1	.063, $p < .802$	.000 (.000-.117)	1.000	.003
Ordstaving	204	2	3.351, $p < .187$	.058 (.000-.162)	.996	.012
Nonordlesing	208	2	4.077, $p < .130$	.071 (.000 - .170)	.995	.015
Nonordstaving	206	2	2.22, $p < .529$	.000 (.000 - .105)	1.000	.017

#### 4.4.2 Bivariate autoregressive modeller

Som nevnt tidligere, er stabiliteten mellom de ulike måletidspunktene høy, og de autoregressive effektene mellom måletidspunktene er sterk. Dette kan skape problemer når man forsøker å modellere kryssladninger. Jeg har derfor testet ut én og én kryssladning med de nærmeste først.

##### Nonordlesing og nonordstaving

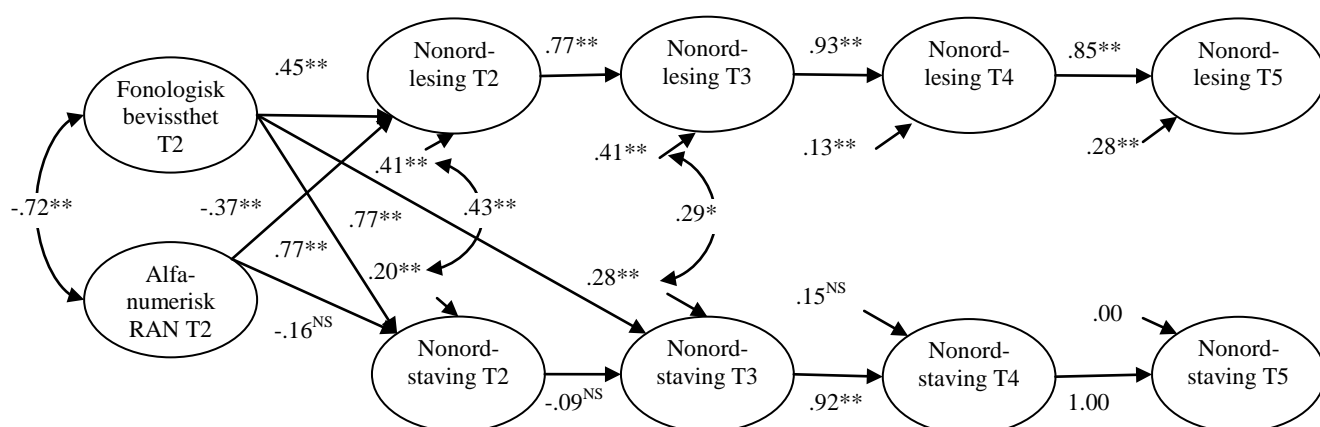
I en bivariat modell med nonordlesing lesing og nonordstaving (figur 15a) er det ingen kryssladninger som er signifikante. Nonordstaving på tidspunkt 5 er satt til 0 på grunn av negativ (men ikke signifikant) residualvarians. Modelltilpasningen er meget god. Det er en sterk sammenheng mellom nonordlesing og nonordstaving på T2 (.79) og en moderat sammenheng mellom residualene til nonordlesing og nonordstaving på T3 (.29). Det er ingen signifikant sammenheng mellom residualene til nonordlesing og nonordstaving på T4 og T5.



**Modelltilpasning:**  $YB\chi^2(17, N = 209) = 12.055, p < .797$ , RMSEA = .000 (90% CI = .000 - .041), CFI = 1.000, SRMR = .024

**Figur 15a** Bivariat latent autoregressiv modell nonordlesing og nonordstaving

Når fonologisk bevissthet og alfanumerisk RAN tas med i modellen, reduseres den opprinnelige sammenhengen mellom nonordlesing og nonordstaving på T2. Fonologisk bevissthet predikerer nonordstaving på både T2 og T3 (.77). Fonologisk bevissthet og alfanumerisk RAN forklarer til sammen 80% av variasjonen i nonordstaving og 59% av variasjonen i nonordlesing på T2. Fonologisk bevissthet og nonordstaving målt på T2 forklarer til sammen 72% av variasjonen i nonordstaving på T3.

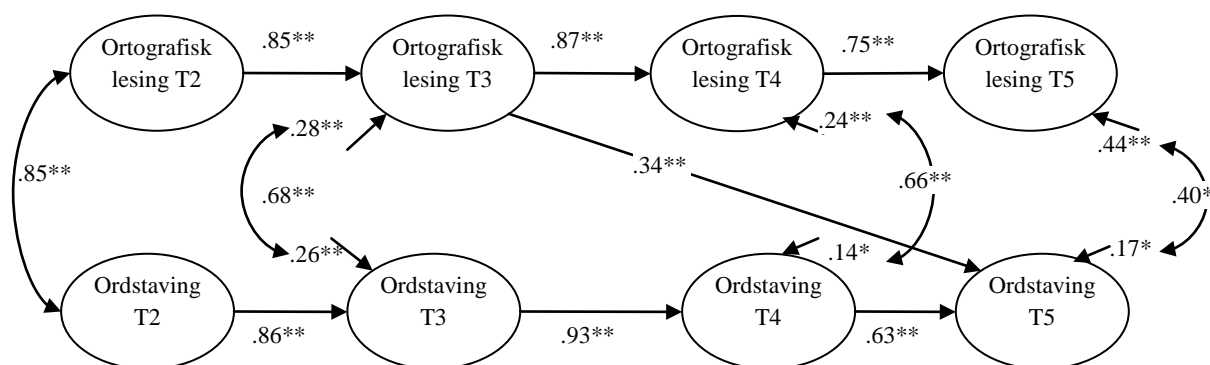


**Modelltilpasning:**  $YB\chi^2(56, N = 209) = 55.372, p < .499$ , RMSEA = .000 (90% CI = .000 - .042), CFI = 1.000, SRMR = .033

**Figur 15b** Bivariat latent autoregressiv modell nonordlesing og nonordstaving med kovariater

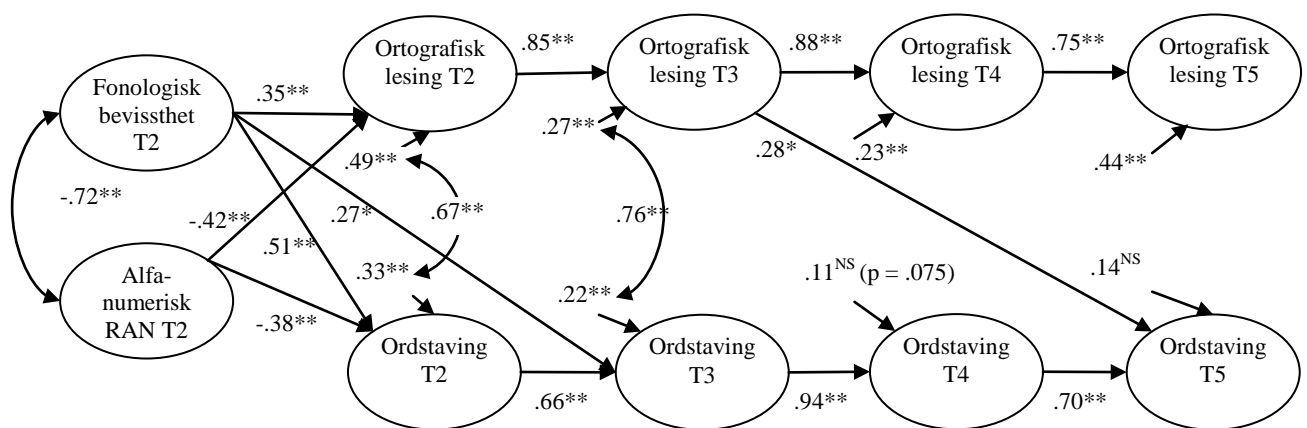
## Ortografisk lesing og ordstaving

Korrelasjonen mellom ortografisk lesing og ordstaving på T2 er sterk (.85). Ortografisk lesing på T3 predikerer dessuten variasjon i ordstaving på T5 (.34).

**Figur 16a** Bivariat latent autoregressiv modell ortografisk lesing og ordstaving

62

modellen uten kovariater, predikerer ortografisk lesing på T3 ordstaving på T5. Alfa-numerisk RAN predikerer variasjon i både ordstaving og ortografisk lesing på T2. Estimaten er på henholdsvis -.39 for ordstaving og -.40 for ortografisk lesing. Fonologisk bevissthet predikerer variasjon i både ordstaving og ortografisk lesing på T2. Estimaten er .43 for ordstaving og .34 for ortografisk lesing. Kovariatene forklarer til sammen 51% av variasjonen i ortografisk lesing og 67% av variasjonen i ordstaving på T2. Fonologisk bevissthet bidrar også til å predikere variasjon i ordstaving på T3 (.29). Alfa-numerisk RAN har større, og fonologisk bevissthet noe mindre forklaringskraft for ordstaving enn for nonordstaving. Som for nonordlesing og nonordstaving er de autoregressive effektene sterke for både ortografisk lesing og ordstaving på alle måletidspunkter.



**Modelltilpasning:**  $YB\chi^2(53, N = 208) = 61.404, p < .200, RMSEA = .028$  (90% CI = .000 - .054), CFI = .994, SRMR = .031

**Figur 16b** Bivariat latent autoregressiv modell ortografisk lesing og ordstaving med kovariater

## 5 Diskusjon

Målet med analysene har vært å undersøke forholdet mellom begynnende lese- og staveutvikling. Er det nonordstaving som forklarer nonordlesing tidlig i utviklingen og ortografisk lesing som forklarer ordstaving litt senere i utviklingen, eller er det slik at lesing og staving utvikler seg sammen og påvirker hverander gjensidig? Disse hypotesene har jeg testet ut både med og uten fonologisk bevissthet og alfanumerisk RAN som kontrollvariabler (kovariater). Et annet mål med oppgaven har vært å drøfte hvor godt de to valgte metodene (autoregressive modeller og latente vekstkurvemodeller) egner seg til å kunne si noe om den gjensidige utviklingen av begynnende lese- og staveferdigheter og forholdet mellom disse.

### 5.1 Lesing og staving av nonord

En oppsummering av resultatene fra analysene av forholdet mellom utviklingen av nonordlesing og nonordstaving viser at ingen av kryssregresjonene er signifikante, verken i den autoregressive modellen eller i den latente vekstkurvemodellen.

I vekstkurvemodellen uten kovariater er det sterk korrelasjon mellom interceptfaktorene til nonordstaving og nonordlesing (.78), men ingen signifikant sammenheng mellom residualene til slopefaktorene. Dette kan tolkes som at *veksten* i nonordlesing og nonordstaving ikke har noe til felles når det kontrolleres for nivået av de respektive ferdighetene på T2. Videre finner jeg at den sterke sammenhengen mellom de to interceptfaktorene i modellene uten kovariater forklares fullt ut av kovariatene fonologisk bevissthet og alfanumerisk RAN.

Ferdigheter i nonordstaving tre måneder etter oppstart av formell lese- og staveopplæring (intercept) forklarer 76% av variasjonen i den lineære veksten (slope). Siden sammenhengen er negativ, betyr det at de barna som er svakest i nonordstaving i starten har den bratteste veksten og til en viss grad tar igjen de som har bedre staveferdigheter på det første måletidspunktet. Videre forklares variasjoner i den kvadratiske faktoren fullt ut av slopefaktoren. Også denne sammenhengen er negativ. Det er altså de barna som har den bratteste veksten i starten som etter hvert har en sterkest avtakende vekst. For nonordlesing finner jeg at det ikke er noen sammenheng mellom hvor man starter og hvordan man vokser.



Resultater fra analysen av de autoregressive modellene viser at sammenhengen mellom lesing og staving av nonord er sterk på T2. Den sterke sammenhengen mellom nonordlesing og nonordstaving på T2 forklares i stor grad av fonologisk bevissthet og alfanumerisk RAN. Fonologisk bevissthet bidrar dessuten sterkt til å predikere variasjoner i staving av nonord på T3 (.77). Denne sammenhengen mellom fonologisk bevissthet på T2 og nonordstaving på T3 fanges ikke opp av den latente vekstkurvemodellen.

Både vekstkurvemodellene og de autoregressive modellene synliggjør at sammenhengen mellom nonordlesing og nonordstaving er sterk på det første måletidspunktet, og denne sammenhengen forklares i stor grad av kovariatene fonologisk bevissthet og alfanumerisk RAN.

Den autoregressive effekten er sterk mellom alle måletidspunkt både for nonordlesing og nonordstaving. Det er altså høy stabilitet i rekkefølgen på barna på alle måletidspunkter; de som er gode i nonordstaving og nonordlesing på T2 er også gode på T3. De som er gode på T3 er gode også på T4, og det samme gjelder mellom T4 og T5. Det er ikke mye varians igjen å forklare når det kontrolleres for effekten av henholdsvis nonordlesing og nonordstaving på tidligere tidspunkt.

Selv om vekstkurvemodellen viser at de barna som var svake på nonordstaving på det første måletidspunktet til en viss grad tar igjen de barna som hadde bedre ferdigheter til å begynne med, viser den autoregressive modellen at det er stor stabilitet i rekkefølgen mellom barna. De som var best i starten av andre klasse er fremdeles best i slutten av fjerde klasse.

## 5.2 Ortografisk lesing og ordstaving

Heller ikke i vekstkurvemodellene med ortografisk lesing og ordstaving finner jeg noen sammenheng mellom ortografisk lesing på starttidspunktet og veksten i ordstaving, eller mellom ordstaving på starttidspunktet og veksten i ortografisk lesing (altså ingen signifikante kryssregresjoner).

I vekstkurvemodellen uten kovariater er det en sterk sammenheng mellom interceptfaktorene og mellom residualene til slopefaktorene. Når kovariatene er med i modellen, er det en høy positiv korrelasjon mellom residualene til interceptfaktorene (.66). Dette reflekterer at ortografisk lesing og ordstaving på starttidspunktet har noe til felles utover det som forklares

av fonologisk bevissthet og alfanumerisk RAN. Det er også en sterk sammenheng mellom residualene til slopefaktorene (.96), og disse residualene er ganske store. Dette betyr at den lineære veksten til ortografisk lesing og ordstaving har noe til felles utover det som forklares av de respektive interceptfaktorene og kovariatene.

Intercept til ordstaving forklarer 35% av variansen i slope, og slope forklarer fullt ut hvordan veksten etter hvert avtar. Negativt fortegn betyr at det er de som starter lavest som vokser mest i starten, og at det er disse som etter hvert også har den sterkeste avtagende veksten. Intercept til ortografisk lesing forklarer 15% av variansen i slope. Her er regresjonen positiv, og tolkes som at de som starter høyest vokser mest til å begynne med. Men denne sammenhengen er ikke veldig sterk. Slope predikerer fullt ut hvordan veksten etter hvert vokser mindre bratt for de som vokste mest i starten (negativt fortegn).

I de autoregressive modellene finner jeg også en sterk sammenheng mellom residualene til ortografisk lesing og ordstaving på T2 og T3 når kovariatene er med i modellen. Residualene på T3 er ikke så store, men man kan allikevel tolke dette som at ortografisk lesing og ordstaving har noe til felles på T2 og T3 utover det som forklares av kovariatene og den autoregressive effekten. Ortografisk lesing på T3 predikerer dessuten variasjoner i ordstaving på T5. Denne sammenhengen fant jeg ikke i vekstkurvemodellen.

Som i modellene med nonordlesing og nonordstaving er den autoregressive effekten sterk mellom alle måletidspunkt både for ortografisk lesing og ordstaving. Det er altså høy stabilitet i rekkefølgen på barna på alle måletidspunkter; de som er best i ortografisk lesing og ordstaving på T2 er også best på T3. De som er best på T3 er best også på T4, og det samme gjelder også mellom T4 og T5. I den autoregressive modellen er det er ikke mye varians igjen å forklare når det kontrolleres for effekten av henholdsvis ortografisk lesing og ordstaving på tidligere tidspunkt.

### 5.3 Forholdet mellom begynnende lese- og staveutvikling

Både Caravolas et al. (2001) og Leppänen et al. (2006) finner i sine studier støtte for Friths (1985) hypotese om at det er stavingen som først forklarer leseutviklingen og så lesingen som forklarer staveutviklingen.

Resultatene av mine analyser gir ingen støtte for at det i den helt tidlige fasen er staving som forklarer leseutviklingen. Jeg finner noe støtte for at det i en senere fase er ortografisk lesing som forklarer ordstavingen. Resultatene fra den autoregressive modellen viser at ortografisk lesing i slutten av andre klasse predikerer staveferdigheter i midten av fjerde klasse. I min undersøkelse er første måletidspunkt i andre klasse, tre måneder etter oppstart av formell lese- og staveopplæring. Norske barn ser ut til å tilegne seg kunnskaper om sammenhengen mellom fonemer og grafemer i løpet av det første og andre året av skolegangen (Lervåg & Hulme, 2010). Ut i fra mine resultater kan det se ut som om utviklingen i forhold til ”den alfabetiske” fasen har kommet et stykke på vei for de fleste allerede tre måneder etter oppstart av formell lese- og staveopplæring. Det er ikke lenger stavingen som forklarer leseutviklingen, men den ortografiske lesingen som forklarer staveutviklingen. Jeg mangler data på lesing og staving i første klasse, og har derfor ikke kunnet teste ut hvordan staving og lesing påvirker hverandre helt tidlig i utviklingen. I tillegg kan det være at mine resultater skiller seg fra Caravolas et al. (2001) fordi engelsk er en mye mer inkonsistent ortografi enn norsk. Når det gjelder resultatene til Leppänen et al. (2006) kan disse blant annet skyldes at de autoregressive effektene i deres studie er mye svakerer enn hos meg. I mine analyser er det rett og slett ikke så mye varians igjen å forklare når det tas hensyn til de autoregressive effektene av lesing og staving på tidligere tidspunkt.

Mine resultater viser videre at fonologisk bevissthet og RAN i stor grad forklarer den tidlige sammenhengen mellom nonordlesing og nonordstaving. Disse kovariatene forklarer også en del av sammenhengen mellom ortografisk lesing og ordstaving på T2, men ikke fullt ut slik som for nonordlesing- og staving.

Ortografisk lesing og ordstaving har altså noe til felles på det første måletidspunktet utover det som forklares av kovariatene. Den videre lineære veksten i ortografisk lesing og ordstaving har også noe felles utover det som forklares av de respektive ferdighetene på første måletidspunkt og kovariatene. Dette kan for eksempel være ferdigheter Ehri (1995, 2000)

omtaler under overskriften ”konsolidert alfabetisk nivå”, der barna lærer at ord består av større enheter og bokstavkonstellasjoner som man finner igjen i ulike ord. Slike større enheter kan være stavelser, deler av stavelser (som opptakter og rim), affikser i begynnelsen eller slutten av ord og morfemer. Dette er ferdigheter som oppnås gjennom praktisering av lesing og staving. Hvor mye det enkelte barn praktiserer lesing og staving, handler ikke bare om det som skjer i klasserommet, men også om forhold som hjemmebakgrunn (for eksempel hvor mye det leses hjemme) og motivasjon.

Mine analyser av ortografisk lesing og ordstaving gir altså noe, om enn ganske svak, støtte til Friths teori om at det er lesingen som forklarer staveutviklingen når barna har nådd et visst nivå når det gjelder kjennskap til sammenhengen mellom grafemer og fonemer. Resultater fra den autoregressive modellen viser at ortografisk lesing målt på slutten av andre klasse predikerer ordstaving i midten av fjerde klasse når det kontrolleres for autoregressive effekter av staving på tidligere tidspunkt.

Vekstkurvemodellen indikerer at det er en gjensidig påvirkning mellom ortografisk lesing og ordstaving hele veien, og gir dermed støtte til Ehris teori. Her er det en sterk sammenheng mellom residualene til de to interceptfaktorene og residualene til de to slopefaktorene. Residualene er ganske store, særlig for de to slopefaktorene. Dette kan som tidligere nevnt, tolkes som at den videre lineære veksten i ortografisk lesing og ordstaving har noe til felles utover det som forklares av de respektive ferdighetene på første måletidspunkt og de to kovariatene.

## 5.4 Latente vekstkurvemodeller

Analyse av univariate latente vekstkurvemodeller synliggjorde at veksten ikke var lineær for noen av variablene. For ortografisk lesing, ordstaving og nonordstaving konkluderte jeg med at en kvadratisk modell beskrev dataene på en god måte. For nonordlesing var det en ”completely latent” modell som så ut til å beskrive dataene best.

Hvilken modell man velger handler ikke bare om god eller dårlig modelltilpasning, men også om tolkbarhet. Et eksempel på dette er de to ulike måtene å representere en ikke-lineær vekst på. Prediksjon av lineær og kvadratisk vekst i en kvadratisk modell forvansker den intuitive

forståelsen av prediksjonen. Tolkbarheten til en ikke-lineær modell der for eksempel de to midterste slopeladningene estimeres fritt (en completely latent modell) kan muligens sies å være mer umiddelbart forståelig. Her tolkes gjennomsnittlig slope som gjennomsnittlig vekst mellom første og siste måletidspunkt. Det er også enklere å regne ut vekstskårer per tidsenhet da veksten mellom to tidspunkter tolkes som en viss prosent av den totale veksten (ved å se på de standardiserte faktorladningene fra slope). På den andre siden mister man informasjon ved å velge en completely latent modell. I min kvadratiske modell av nonordstaving er det de som vokser mest i starten som etter hvert har en sterkest avtagende vekst. Denne informasjonen går man glipp av i en completely latent modell.

Når utviklingen er ikke-lineær, vil individuelle forskjeller variere som et resultat av at det er forskjeller i vekstspurter og plataår. I slike tilfeller er det viktig at man samler inn data lenge nok og tett nok for å kunne skille mellom en midlertidig forbedring og en mer permanent reduksjon av individuelle forskjeller (Parilla et al. 2005). Mine resultater, der jeg finner at veksten er ikke-lineær for alle de fire lese- og stavevariablene, synliggjør viktigheten av å måle lenge nok når man skal si noe om utviklingen av denne typen ferdigheter. Dersom man kun hadde målt på T2 og T3, ville man ha fått et ufullstendig bilde av utviklingen. Da hadde man konkludert med at de svakeste hadde brattest vekst, og gått glipp av informasjon som indikerer at denne veksten er kortvarig.

Til tross for de mange fordelene med latente vekstkurvemodeller, er det ikke alltid at denne analysemetoden er den mest hensiktsmessige. En av de viktigste innvendingene er at metoden krever relativt store utvalg.

I min undersøkelse ser det ikke ut som de latente vekstkurvemodellene fanger opp nyansene i relasjonen mellom lesing og staving eller den mer tidsspesifikke effekten av fonologisk bevissthet målt i starten av andre klasse på nonordstaving i slutten av andre klasse.

## 5.5 Autoregressive modeller

Analysen av univariate latente autoregressive modeller viser sterke autoregressive effekter mellom alle måletidspunkter for alle de fire lese- og stavevariablene. Dette innebærer at det er lite variasjon igjen å forklare.

Den autoregressive modellen har blitt kritisert for kun å reflektere forandringer i hele gruppen. Stoolmiller & Banks (1995) poengterer en annen svakhet ved autoregressive modeller, nemlig at noen utviklingsprosesser kan være stabile over tid og at det dermed er vanskelig å vise at utviklingen er korrelert med en tredjevariabel. De sier videre at at utvikling kan skje som en følge av forandringer i andre kovariater, men dersom denne forandringen ikke involverer endring i rekkefølgen mellom individene eller i formen til distribusjonen av den avhengige variabelen, så vil forandringen i stor grad tilskrives den autoregressive effekten.

Stoolmiller & Banks (1995) poengterer at mengden varians (individuelle forskjeller) i endring over tid kan spille en avgjørende rolle for å bestemme om det er en autoregressiv modell eller andre modeller som klarer å synliggjøre en effekt av et kovariat på denne endringen. Ulike modeller kan altså komme til forskjellige konklusjoner om viktigheten av prediktorer for utvikling. Når variansen (altså individuelle forskjeller) på senere måletidspunkter er stabile eller avtakende i forhold til variansen på det første måletidspunktet, vil den autoregressive modellen være nyttigst til å avdekke mulige prediktorer for utviklingen, mens andre modeller vil være nyttige når variansen (individuelle forskjeller) på senere måletidspunkter er store i forhold til variansen på det første måletidspunktet.

I min undersøkelse ser det ut til at de autoregressive modellene er bedre egnet enn de latente vekstkurvmodellene til å fange opp nyansene i relasjonen mellom lesing og staving. Dersom Friths (1985) teori stemmer, er det snakk om tidsspesifikke relasjoner, ikke relasjoner der veksten er generalisert over alle måletidspunktene. Som nevnt tidligere, predikerer ortografisk lesing på slutten av andre klasse ordstaving i midten av fjerde klasse i den autoregressive modellen. Denne sammenhengen fanges ikke opp av vekstkurvmodellen. I vekstkurvmodellen fanges heller ikke sammenhengen mellom fonologisk bevissthet og nonordstaving på T3 opp. Dette fordi fonologisk bevissthet har størst betydning for nonordstaving helt i starten av utviklingen. Denne sammenhengen forsvinner i en vekstkurvmodell der veksten generaliseres over så mange år som den gjør i min

undersøkelse. Det ser altså ut til at de autoregressive modellene fanger opp tidsspesifikke relasjoner bedre enn vekstkurvemodellene.

## 5.6 Alternative modeller

De to metodene jeg har benyttet meg av har ofte blitt sett på som ”konkurrenter” til hverandre. Den autoregressive modellen har hovedfokus på ”fixed effects”, altså gjennomsnitt, mens latente vekstkurvemodeller har hovedfokus på ”random effects”, altså individuelle forskjeller i vekstkurver over tid. Bollen & Curran (2004) argumenterer for en syntese av disse to tradisjonene fordi det finnes tilfeller i samfunnsvitenskapen der både autoregressive prosesser og individuelle vekstkurveprosesser pågår samtidig. En slik modell kalles en ALT-modell (autoregressive latent trajectory model). Som for alle andre modeller innenfor SEM-tradisjonen må en slik modell begrunnes substansielt. Slike kompliserte modeller kan også være vanskeligere å tolke.

Et kritisk spørsmål når man studerer vekst, er hvorvidt alle individer følger samme vekstkurve. Parilla et al. (2005) tar opp problemstillingen om variasjoner i vekstkurvekomponentene skyldes flere distinkte utviklingskurver. Dette kan være tilfelle dersom utvalget er heterogent og inkluderer forskjellige undergrupper som følger ulike vekstmønstre. For eksempel kan man i et utvalg finne en gruppe med lave startferdigheter og påfølgende sakte vekst, en gruppe med høyere startferdigheter og høyere påfølgende vekst og en gruppe med lave startferdigheter og påfølgende rask vekst. Analysemetoden Growth Mixture Modeling adresserer dette ved å identifisere latente klasser. Dersom man finner slike ulike vekstmønstre, er det et interessant å se om man kan finne ferdigheter som kan predikere hvilken vekstkurve barna mest sannsynlig kommer til å følge. Lervåg & Hulme (2010) finner gjennom slike analyser at veksten i ordstaving best kunne karakteriseres som variasjoner rundt én vekstkurve, mens veksten i nonordstaving kunne karakteriseres som variasjoner rundt to ulike vekstkurver. En mulig utvidelse av mine analyser kunne ha vært å se om det hadde vært mulig å identifisere ulike vekstkurver for de to lesevariablene ortografisk lesing og nonordlesing.

## 5.7 Hvorfor ikke se seg blind på modelltilpasningen

I statistiske analysemetoder som autoregressive modeller og vekstkurvemodeller, kan det være en fare at man ser seg blind på ”enkle” statistiske mål som overordnet modelltilpasning. Et eksempel fra mine data som kan illustrere dette, er mine vekstkurveanalyser av nonordlesing og nonordstaving. Ved kun å se på de ulike modelltilpasningsmålene kan man få inntrykk av at det er en modell der alfanumerisk RAN er med som eneste kovariat som er den “beste” modellen, og ihvertfall bedre enn en modell der kun fonologisk bevissthet er med som kovariat. Når man går videre og ser på parameterestimaterne og forklart varians i de latente vekstfaktorene, blir bildet et annet. Forklart varians i interceptfaktorene er størst når begge kovariatene er inne i modellen, selv om den overordnede modelltilpasningen er vel så god når alfanumerisk RAN er inne i modellen som eneste kovariat. Alfa-numerisk RAN mister dessuten sin i utgangspunktet sterke forklaringskraft for nonordstaving når det kontrolleres for effekten av fonologisk bevissthet. Teoretisk sett gir det heller ikke mening med en modell der kun RAN er med som kovariat.

Selv om SEM er en fleksibel statistisk analysemetode, er det viktig å være klar over at

“SEM is not, however, a statistical magic bullet. It cannot be used to prove that a model is correct and it can not compensate for a poorly designed study. In addition, even a well-fitting SEM model can have problematic lower-order components and omit important variables” (Tomarken & Waller, 2005, s. 56).

## 5.8 Avsluttende kommentarer

I de autoregressive modellene var stabiliteten høy for alle de fire lese- og stavevariablene i hele perioden som ble målt. Dette innebærer at det er lite variasjon igjen å forklare. Dette kan være årsaken til at jeg ikke fant flere signifikante kryssladninger mellom lesing og staving i de autoregressive modellene.

Når det gjelder vekstkurvemodellene, er det mye uforklart varians i slopefaktorene, også når det kontrolleres for effekten av fonologisk bevissthet og alfanumerisk RAN og de respektive lese- og stavevariablene målt på starttidspunktet. Dermed er det ikke manglende varians som gjør at kryssregresjonene ikke blir signifikante. Jeg spør om det kan være at veksten er



generalisert over for lang tid til at modellen klarer å fange opp en eventuell direkte påvirkning fra ortografisk lesing ved første måletidspunkt på den videre veksten av ordstaving.

En svakhet ved de dataene jeg har analysert, er manglende målinger av lesing og staving i første klasse.

Vekstkurvemodellene gir mer detaljert informasjon om vekst enn de autoregressive modellene ved at de sier noe om sammenhengen mellom hvor man starter og hvordan man vokser. De autoregressive modellene ser ut til å fange opp tidspesifikke relasjoner noe bedre.

De statistiske analysemetodene innenfor SEM-familien utvikles stadig og blir mer og mer sofistikerte og fleksible med mulighet for mer kompleks modellering. Det er alltid en fare for at denne kompleksiteten går ut over tolkbarheten. Nok en gang er det derfor viktig å poengtere at modeller alltid må begrunnes substansielt, og at man ikke må seg seg blind på overordnet modelltilpasning eller forslag i modifikasjonsindeksen. Det er uansett umulig å bevise at en modell er korrekt gjennom statistiske analyser. Alternative modeller kan passe like godt eller bedre til dataene.



# Litteraturliste

- Bast, J., & Reitsma, P. (1997). Mathew Effects in Reading: A Comparison of Latent Growth Curve Models and Simplex Models with Structured Means. *Multivariate Behavioral Research*, 32(2), 135-167.
- Bollen, K. A., & Curran, P. J. (2004). Autoregressive Latent Trajectory (ALT) Models. A Synthesis of Two Traditions. *Sociological Methods & Research*, 32(3), 336-383.
- Bollen, K. A., & Curran, P. J. (2006). *Latent Curve Models. A Structural Equation Perspective*. Hoboken NJ: Wiley.
- Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press.
- Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The Foundations of Spelling Ability: Evidence from a 3-Year Longitudinal Study. *Journal of Memory and Language*, 45, 751-774.
- de Jong, P. F., & van der Leij, A. (1999). Specific Contributions of Phonological Abilities to Early Reading Acquisition: Results from a Dutch Latent Variable Longitudinal Study. *Journal of Educational Psychology*, 91(3), 450-476.
- Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues and Applications*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Ehri, L. C. (1995). Phases of Development in Learning to Read Words By Sight. *Journal of Research in Reading*, 18(2), 116-125.
- Ehri, L. C. (2000). Learning To Read and Learning To Spell: Two Sides of a Coin. *Topics in Language Disorders*, 20(3), 19-36.

- Fitzgerald, J., & Shanahan, T. (2000). Reading and Writing Relations and Their Development. *Educational Psychologist*, 35(1), 39-50.
- Frith, U. (1985). Beneath the Surface of Developmental Dyslexia. I K. E. Patterson, J. C. Marshall, & M. Coltheart (Red.), *Surface Dyslexia: Neuropsychological and Cognitive Studies of Phonological Reading* (s. 301-330). London: Lawrence Erlbaum Associates.
- Hagtvet, K. A. (2010). Strukturelle likningsmodeller: Målingsmodell og regresjonsmodell. I M. Martinussen (Red.), *Kvantitativ forskningsmetodologi i samfunns- og helsefag* (s. 285-314). Bergen: Fagbokforlaget.
- Hulme, C., & Snowling M. J. (2009). *Developmental Disorders of Language Learning and Cognition*. UK: Wiley-Blackwell.
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff Criteria For Fit Indexes In Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. New York: The Guilford Press.
- Landerl, K., & Wimmer, H. (2008). Development of Word Reading Fluency and Spelling in a Consistent Orthography: An 8-Year Follow-Up. *Journal of Educational Psychology*, 100(1), 150-161.
- Leppänen, U., Nieme, P., Aunola, K., & Nurmi, J.-E. (2004). Development of Reading Skills Among Preschool and Primary School Pupils. *Reading Research Quarterly*, 39, 72-93
- Leppänen, U., Nieme, P., Aunola, K., & Nurmi, J.E. (2006). Development of Reading and Spelling Finnish From Preschool to Grade 1 and Grade 2. *Scientific Studies of Reading*, 10(1), 3-30.

- Lervåg, A. (2005). *Prediction of development in beginning reading and spelling. A Norwegian latent variable study*. Akademisk avhandling. Universitetet i Oslo.
- Lervåg, A., Bråten, I., & Hulme, C. (2009). The Cognitive and Linguistic Foundations of Early Reading Development: A Norwegian Latent Variable Longitudinal Study. *Developmental Psychology*, 45(3), 764-781.
- Lervåg, A., & Hulme, C. (2009). Rapid Automatized Naming (RAN) Taps a Mechanism That Places Constraints on the Development of Early Reading Fluency. *Psychological Science*, 20(8), 1040-1048.
- Lervåg, A., & Hulme, C. (2010). Predicting the Growth of Early Spelling Skills: Are There Heterogeneous Developmental Trajectories? *Scientific Studies of Reading*, 14(6), 1-29.
- Logan, J. A. R., Schatschneider, C., & Wagner, R. (2011). Rapid serial naming and reading ability: the role of lexical access. *Reading and Writing*, 24(1), 1-25.
- Lyster, S-A. H. (1998). *Å lære å lese og skrive - Individ i kontekst*. Oslo: Gyldendal Akademisk
- Muthén, L.K., & Muthén, B.O. (1998-2010). *Mplus User's Guide. Sixth Edition*. Los Angeles, CA: Muthén & Muthén.
- Parrila, K. R., Aunola, K., Leskinen, E., Nurmi, J.E., & Kirby, J.R. (2005). Development of Individual Differences in Reading: Results From Longitudinal Studies in English and Finnish. *Journal of Educational Psychology*, 97(3), 299-319.
- Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N.E. (2008). *Latent Growth Curve Modeling*. Los Angeles, CA: Sage Publications.

- Schlueter, E., Davidov, E., & Schmidt, P. (2007). Applying Autoregressive Cross-Lagged and Latent Growth Curve Models to a Three-Wave Panel Study. I K. Montford, H. Oud, & A. Satorra (Red.), *Longitudinal Models in the Behavioral and Related Sciences* (s. 315-336). Mahwah, NJ: Lawrence Erlbaum Associates.
- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation Literacy Acquisition in European Orthographies. *British Journal of Psychology*, 94, 143-174.
- Seymour, P. H. (2005). Early Reading Development in European Orthographies. I M. J. Snowling, & C. Hulme (Red.), *The Science of Reading. A handbook* (s. 296-315). Oxford UK, Blackwell Publishing Ltd.
- Stoolmiller, M., & Bank, L. (1995). Autoregressive Effects in Structural Equation Models: We See Some Problems. I J. M. Gottman (Red.), *The Analysis of Change* (s. 261-276). Mahwah, NJ: Erlbaum.
- Tomarken, A. J., & Waller, N. G. (2005). Structural Equation Modeling: Strengths, Limitations, and Misconceptions. *Annual Review of Clinical Psychology*, 1, s. 31-65.
- Torgesen, J.K., Wagner, R.K., Rashotte, C. A., Burgess, S., & Hecht, S. (1997). Contributions of Phonological Awareness and Rapid Automatic Naming Ability to the Growth of Word-Reading Skills in Second- to Fifth-Grade Children. *Scientific Studies of Reading*, 1(2), 161-185.
- Verhagen, W., Aarnoutse, C., & van Leeuwe, J. (2008). Phonological Awareness and Naming Speed in the Prediction of Dutch Children's Word Recognition. *Scientific Studies of Reading*, 12(4), 301-324.
- Willett, J. B. (1997). Measuring Change: What Individual Growth Modeling Buys You. I E. Amsel & K. A. Renninger (Red.), *Change and Development: Issues of Theory, Method, and Application*. Mahwah, NJ: Lawrence Erlbaum Associates, s. 213-243.

# Vedlegg

# Vedlegg 1 Korrelasjoner mellom de observerte variablene

	Fe_2	Fs_2	Fi_2	Ran_n_2	Ran_l_2	Non-les2	Non-les3	Non-les4	Non-les5	Ort-les2	Ort-les3	Ort-les4	Ort-les5	Non-stav2	Non-stav3	Non-stav4	Non-stav5	Stav2	Stav3	Stav4	Stav5
Fe_2	1.000																				
Fs_2	.46**	1.000																			
Fi_2	.56**	.49**	1.000																		
Ran_n_2	-.34**	-.34**	-.46**	1.000																	
Ran_l_2	-.40**	-.39**	-.55**	.72**	1.000																
Nonles2	.52**	.40**	.46**	-.53**	-.57**	1.000															
Nonles3	.35**	.20**	.43**	-.48**	-.52**	.68**	1.000														
Nonles4	.28**	.22**	.42**	-.44**	-.48**	.63**	.86**	1.000													
Nonles5	.26**	.14 <sup>NS</sup>	.31**	-.35**	-.35**	.49**	.76**	.79**	1.000												
Ortles2	.45**	.21**	.38**	-.36**	-.42**	.66**	.53**	.47**	.42**	1.000											
Ortles3	.38**	.22**	.42**	-.45**	-.44**	.60**	.65**	.58**	.54**	.57**	1.000										
Ortles4	.35**	.22**	.33**	-.35**	-.38**	.51**	.65**	.58**	.52**	.51**	.73**	1.000									
Ortles5	.29**	.18*	.23**	-.28**	-.25**	.39**	.47**	.50**	.57**	.43**	.60**	.67**	1.000								
Nonstav2	.53**	.52**	.64**	-.53**	-.57**	.68**	.48**	.46**	.35**	.49**	.55**	.46**	.38**	1.000							
Nonstav3	.43**	.39**	.56**	-.46**	-.53**	.48**	.43**	.45**	.36**	.32**	.42**	.39**	.36**	.56**	1.000						
Nonstav4	.36**	.40**	.44**	-.28**	-.38**	.40**	.34**	.36**	.28**	.35**	.35**	.30**	.29**	.43**	.55**	1.000					
Nonstav5	.30**	.22**	.36**	-.32**	-.38**	.38**	.39**	.39**	.34**	.37**	.36**	.30**	.24**	.45**	.47**	.48**	1.000				
Stav2	.47**	.42**	.59**	-.53**	-.63**	.67**	.54**	.54**	.41**	.60**	.62**	.48**	.40**	.73**	.55**	.49**	.45**	1.000			
Stav3	.46**	.36**	.55**	-.45**	-.46**	.56**	.53**	.51**	.42**	.47**	.65**	.57**	.48**	.64**	.59**	.48**	.44**	.71**	1.000		
Stav4	.45**	.33**	.51**	-.37	-.41**	.59**	.57**	.53**	.44**	.50**	.62**	.64**	.51**	.61**	.53**	.49**	.42**	.63**	.75**	1.000	
Stav5	.45**	.31**	.55**	-.45	-.49**	.61**	.62**	.63**	.58**	.48**	.68**	.63**	.60**	.63**	.63**	.49**	.53**	.63**	.67**	.71**	1.000

Note. \*\* p < 0.01

\* p < 0.05